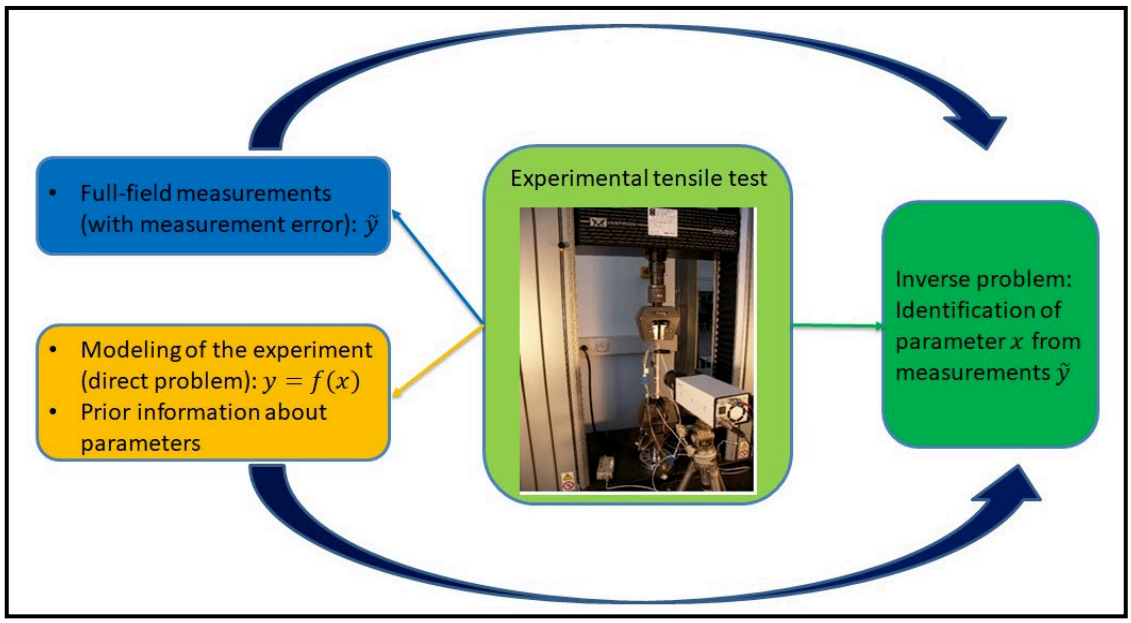
Interval uncertainty method to treat inconsistent measurements in inverse problems

Krushna Shinde

Par **Krushna SHINDE**

Interval uncertainty method to treat inconsistent measurements in inverse problems

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC



Soutenue le 24 février 2021
Spécialité : Mécanique Numérique : Unité de recherche en Mécanique - Laboratoire Roberval (FRE UTC - CNRS 2012)
D2594

Interval uncertainty method to treat inconsistent measurements in inverse problems

Krushna SHINDE

Spécialité : Mécanique Numérique

Thesis defended on 24/02/2021, before the jury composed of:

Jury members:

Reviewers:
GALICHET Sylvie, Professeur des universités
Informatique, LISTIC - Polytech Annecy-Chambéry

CORET Michel, Professeur des universités
Mécanique, Mécanique des Matériaux, GeM, École Centrale de Nantes

Examiners:
PUEL Guillaume, Professeur des universités
Mécanique numérique, MSSMAT, Centrale Supélec

DREVELLE Vincent, Maître de conférences
Informatique, RAINBOW, Université de Rennes 1, IRISA, INRIA Rennes – Bretagne Atlantique, Campus Universitaire de Beaulieu

DRUESNE Frédéric, ECC-HDR
Laboratoire Roberval, Université de Technologie de Compiègne

Supervisors:
FEISSEL Pierre, Professeur des Universités - HDR, Laboratoire Roberval
DESTERCKE Sébastien, Chargé de Recherche au CNRS,Heudiasyc

Université de Technologie de Compiègne
Laboratoire Roberval de Mécanique UMR-CNRS 7337

To my family. Loving mother, supportive father, driven brother and lovely Compiègne

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors Prof. Pierre Feissel and Sebastien Destercke (CNRS researcher in computer science), for the continuous support of my Ph.D. study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis.

Besides my supervisors, I would like to thank the Roberval Laboratory's digital mechanics team for conducting a scientific day every year to discuss research activities within the team.

I thank my colleagues from Roberval in Compiegne, with whom I spent a huge amount of time during office hours and beyond. Thanks to all of you for the moments we shared.

I wish to thank the unconditional support of my parents and my brother despite the long distance. Without their support, it would not have been possible to pursue a doctorate in France.

Finally, I would like to gratefully acknowledge the support provided for this work under the framework of the Labex MS2T, financed by the French Government through the program Investments for the future managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02).

Abstract

The inverse problem consists of finding the parameter values of a physical model given a set of measurements. In mechanical engineering problems, material behavior's characterization requires an inverse method to identify the material parameters. The deterministic identification problem is generally sensitive to data, and one way to resolve this issue is to consider uncertainties in the data. While several such methods exist in the literature, most of them use least-square minimization or Bayesian approaches. This thesis explores how non-probabilistic uncertainty (interval-based) approaches can help obtain a solution to the inverse problem, particularly when measurements are inconsistent with one another. Our approach intrinsically differs from the previously mentioned ones. It does not rely on minimizing the average error but rather on selecting a subset of consistent measurements.

The identification strategy is based on the set theory developed, which allowed us to take into account both prior information about the parameters and measurement uncertainty in the form of sets (interval or boxes) during the inversion process. In this strategy, we developed some indicators of consistency of the measurements to characterize inconsistent measurements, i.e., outliers in the data. We applied this strategy to identify the elastic parameters of an isotropic material. The main advantage of this strategy is that it helps to obtain a feasible set of parameters, but that it can also detect the outliers among noisy measurements. The strategy is subsequently combined with surrogate modeling for identification problems in high dimensional settings. We also applied our strategy to detect damage in the material.

Keywords: Inverse problem, Uncertainty representation, Interval method, Outliers.

Résumé

Le problème inverse consiste à trouver les valeurs des paramètres d'un modèle physique à partir d'un ensemble de mesures. Dans les problèmes de génie mécanique, la caractérisation du comportement des matériaux nécessite une méthode inverse pour identifier les paramètres des matériaux. Le problème d'identification déterministe est généralement sensible aux données, et une façon de résoudre ce problème est de prendre en compte les incertitudes dans les données. Bien que plusieurs de ces méthodes existent dans la littérature, la plupart d'entre elles utilisent la minimisation des moindres carrés ou les approches bayésiennes. Cette thèse explore comment les approches d'incertitude non probabiliste (basées sur des intervalles) peuvent aider à obtenir une solution au problème inverse, en particulier lorsque les mesures sont incompatibles les unes avec les autres. Notre approche diffère intrinsèquement de celles mentionnées précédemment. Il ne repose pas sur la minimisation de l'erreur moyenne mais plutôt sur la sélection d'un sous-ensemble de mesures cohérentes.

La stratégie d'identification est basée sur la théorie des ensembles développée, qui nous a permis de prendre en compte à la fois les informations préalables sur les paramètres et l'incertitude de mesure sous forme d'ensembles (intervalle ou boîtes) lors du processus d'inversion. Dans cette stratégie, nous avons développé des indicateurs de cohérence des mesures pour caractériser les mesures incohérentes, c'est-à-dire les valeurs aberrantes dans les données. Nous avons appliqué cette stratégie pour identifier les paramètres élastiques d'un matériau isotrope. Le principal avantage de cette stratégie est qu'elle permet d'obtenir un ensemble réalisable de paramètres, mais qu'elle peut également détecter les valeurs aberrantes parmi les mesures bruyantes. La stratégie est ensuite combinée à une modélisation de substitution pour les problèmes d'identification dans des environnements de grande dimension. Nous avons également appliqué notre stratégie pour détecter les dommages dans le matériau.

Mots clés: Problème inverse, Représentation de l'incertitude, Méthode d'intervalle, Valeurs aberrantes.

Contents

List of figures

List of tables

Introduction

Material characterization is one of the key element for engineers to design the optimum structures which offer the required performance. The characterization of material behavior requires (under heterogeneous conditions, e.g., complex loading, geometry, or material) an inverse method to identify the material parameters. Identifying the parameters of a physical model from a set of measurements is a common task in many fields such as image processing (tomographic reconstruction [Chetih and Messali, 2015]), acoustic(source identification [Eller and Valdivia, 2009]), or mechanics (material parameter identification [Tam et al., 2017]). Such a problem is known as the inverse problem and is the converse of the so-called forward problem, that simply consists in estimating output values from the model, knowing its parameters and input values. While the forward problem is usually well-posed, it is not the case of the inverse problem [Hadamard, 1902]. Indeed, whenever there is noise in the measurements or error in the model, such a problem may well end-up having no solutions [Tikhonov, 1995] as the number of measurements is typically quite higher than the number of parameters to retrieve. The deterministic identification problem is generally highly sensitive to data quality, and one way to resolve this issue is to take into account uncertainties in the data.

Common solutions to this issue that have been proposed in the literature are to consider either Least-square minimization techniques [Teughels and Roeck, 2005] or Bayesian approaches [Gogu et al., 2010] to cope with the noise in measurements. The first approach simply tries to find the parameter values minimizing the average distance (i.e., the squared error) between the reconstructed and the observed measurements. While this approach may be very efficient (especially for linear models [Kong et al., 2019]). A way around it is the use of Bayesian approaches, where measurement and model errors are modeled by probability distributions, and where a posterior distribution over parameter values is obtained by combining these distributions with a prior distribution on the parameters.

Both these approaches, however, can be quite sensitive to outliers [Blais, 2010, Chen et al., 2000] or to aberrant measurements. In addition to that, Bayesian approaches require to carefully model the errors and to assess precise probabilities, which may not be easy to get in some situations. Indeed, many authors have

argued that in case of incomplete information about the noise, precise Bayesian approaches often require to make extra assumptions that may be false, hence may lead to misleading results [Chen et al., 2000]. As an alternative, there are other non-probabilistic approaches, such as imprecise probabilities [Walley, 1991], intervals [Moore, 1966], fuzzy sets [Zadeh, 1965], possibility theory [Baudrit and Dubois, 2006], and theory of belief function (evidence theory [Shafer, 1976]), to model the uncertainty on the imprecise information.

In the context of the inverse problem, non-probabilistic methods is a less explored area. In this work, we explore how interval-based approaches can be used to obtain a solution to the inverse problem, in particular when measurements are inconsistent with one another. Interval-valued approaches [Moens and Hanss, 2011, Ferson and Ginzburg, 1996, Helton and Johnson, 2011] make a minimal amount of assumptions about the nature of the involved uncertainties (epistemic or aleatory) and only require to define the region in which should be the measurement. In the context of the interval-valued approaches, there are two main challenges: (1) representing prior information about parameters and measurements, and (2) determining the consistency of measurements. The present work addresses these issues by proposing a set-valued inverse method that is not only able to identify a feasible set of the parameters but is also able to detect inconsistencies and outliers in the measurements. Our approach uses intervals or sets [Jaulin et al., 2001] to model uncertainty on the information. Recently [Faes et al., 2019, Mierlo et al., 2019, Fedele et al., 2012, Jaulin and Walter, 1993, Braems et al., 2001], a few interval-based methods have been proposed to identify parameters from measurements. These methods do not specifically address the issues of determining the inconsistency of measurements, which is a one of the goal of this work.

In this work, we propose an inverse strategy relying on interval analysis to deal with uncertain measurements and methods to detect inconsistent measurements (outliers). We apply the proposed strategy in experimentation concerning the identification of material elastic parameters in the presence of possibly inconsistent measurements (here, full-field displacements [Peters and Ranson, 1982]). In the context of structural damage detection, we apply the proposed strategy to show how available measurement data is useful to estimate the location of the damage in the material. The thesis is organized as follows.

- Chapter 1 presents the methods to represent uncertainty on the information, including probability theory and interval theory. It also introduces the inverse methods with probabilistic and non- probabilistic uncertainty representation. Concerning the non-probabilistic approaches, we will limit ourselves to the interval-valued case as our work focus on this particular case and extensions

to other theories is left for future investigations.

- Chapter 2 presents the basic identification strategy based on intervals and sets. It also discusses our proposed outlier detection algorithms and numerical implementation of the identification strategy.

- Chapter 3 presents the application of the identification strategy to the mechanical inverse problem and how to deal with inconsistent measurements in inverse problems. It also compares of outlier detection criteria among themselves as well as the proposed identification strategy with the Bayesian inference method in terms of sensitivity to outliers.

- Any identification strategy becomes computationally complex when either a large number of parameters to identify or available measurements are in the large number. Chapter 4 presents how surrogate modeling can be useful when the identification strategy used with high dimensional data. It also illustrates one more critical application of identification strategy, in the context of structural damage detection, to estimate the location of the damage in the material.

Inverse problem and uncertainty representation methods

Contents

1.1 Inverse problem

We use inverse problems for predicting or characterizing the parameters of any physical model from experimental measurements given the physics of the model under study. The inverse problem is the reverse of a forward or direct problem. Forward problems deal with computing the response of any physical model knowing its parameters. It can be stated in a simple example as:

$$\text{find } \boldsymbol{u} = \boldsymbol{u}(\boldsymbol{\theta}) \text{ such that } \boldsymbol{g}(\boldsymbol{u}, \boldsymbol{\theta}) = \mathbf{0}, \text{with } \boldsymbol{\theta} \text{ given} \tag{1.1}$$

where $\boldsymbol{u}$ is the response of the model, $\boldsymbol{\theta}$ denotes the model parameters, and $\boldsymbol{g}$ is the function derived from the physics. The inverse problem can be formulated as follows:

$$\text{find } \boldsymbol{\theta} \in \boldsymbol{\Theta} \text{ such that } \boldsymbol{g}(\tilde{\boldsymbol{u}}, \boldsymbol{\theta}) = \mathbf{0}, \tag{1.2}$$

where $\tilde{\boldsymbol{u}}$ is the measured response corresponding to $\boldsymbol{u}$, and $\boldsymbol{\Theta}$ denotes the parameter space in which $\boldsymbol{\theta}$ is searched. Unlike standard forward problems, inverse problems are usually ill-posed because they do not satisfy the Hadamard well-posedness conditions [Hadamard, 1902]: (1) a solution exists; (2) the solution is unique; (3) the solution's behaviour changes continuously with the initial conditions. Both the forward and inverse problems make use of information. In the non-deterministic case, uncertainty on the information plays a significant role in the solution of the forward and inverse problems. For the first, the uncertainty on input parameter information, i.e. parameter uncertainty and uncertainty on the numerical model which approximates the physical model,i.e. the model uncertainty affects the solution. For the inverse problem, in addition to the parameter and model uncertainty, uncertainty on the measurement information affects the solution. Modeling uncertainty in the information can be useful to formulate a well-posed problem.

Uncertainty is inevitable in the inverse problem. Many researchers divide it into two types [Hora, 1996, Helton et al., 2004]: aleatoric uncertainty arises due to the system or physical quantity variability; epistemic uncertainty arises due to the lack of knowledge or incomplete information. The first one is irreducible, has a random nature, and is generally represented by a probability distribution when sufficient information is available. While the second one, in contrast to aleatoric uncertainty, is reducible with additional information. The experimental information related to data measured during the experimental tests, i.e., measurement uncertainty, is often linked to aleatory uncertainty. In contrast, the model and parameter uncertainties are often linked to epistemic uncertainty. The use of a probabilistic approach for

the representation of epistemic uncertainty often raised questions. In the literature, a lot of researchers argued that probabilistic methods are not well suited for representing and propagating epistemic uncertainty [Moens and Hanss, 2011, Ferson and Ginzburg, 1996, Helton and Johnson, 2011]. As an alternative, there are non-probabilistic approaches such as imprecise probabilities [Walley, 1991], intervals [Moore, 1966], fuzzy sets [Zadeh, 1965], possibility theory [Baudrit and Dubois, 2006], and theory of belief function (evidence theory [Shafer, 1976]) to model the uncertainty of incomplete or imprecise information. As our thesis work focuses on non-probabilistic uncertainty representation with intervals and sets, we will limit ourselves to interval theory. In this thesis work, we emphasize on how to deal with inconsistent measurements in the inverse problem. In the next section, we review existing deterministic and non-deterministic inverse methods.

1.2 Deterministic inverse problem

In the context of the deterministic inverse problem, a way to solve the inverse problem in the presence of measurement uncertainty is to find the value $\boldsymbol{\theta}$ minimizing the error between the reconstructed data and the observed measurements. If one adopts a least-square minimization, this comes down to solve the optimization problem as

$$\boldsymbol{\theta}^{opt} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\text{Arg Min}}\ J(\boldsymbol{\theta}) \tag{1.3}$$

where $J(\boldsymbol{\theta}) = \|\boldsymbol{u}(\boldsymbol{\theta}) - \tilde{\boldsymbol{u}}\|_2^2$. Non-uniqueness and ill-posedness of the solution are inevitable in the inverse problem. To solve this problem, some authors suggested regularization [Engl and Ramlau, 2015] in the optimization problem. The idea is to explore prior information about parameters in addition to the measurement information. The Tikhonov regularization [Tikhonov et al., 1995] technique takes prior information about parameters into account by adding a stabilizing function to the original objective function as

$$\mathcal{J}(\boldsymbol{\theta}) = J(\boldsymbol{\theta}) + \alpha\mathcal{R}(\boldsymbol{\theta}), \tag{1.4}$$

where $J(.)$ defines the discrepancy between the predicted data and measurements. The coefficient α is the regularization parameter, and $\mathcal{R}(.)$ is the stabilizing function. One of the ways to choose $\mathcal{R}(.)$ is

$$\mathcal{R}(\boldsymbol{\theta}) = \|\boldsymbol{\theta} - \boldsymbol{\theta_0}\|_2^2, \tag{1.5}$$

where $\boldsymbol{\theta}$ is considered to be close to the prior value $\boldsymbol{\theta_0}$. The selection of the coefficient α is a vital point for Tikhonov regularization. There are various ways to select the optimal value of α such as the L-curve method [Hansen, 1992, Kindermann and Raik, 2020].

The identification method described above is a part of deterministic strategies. They mainly focus on predicting the parameters that best fit with the measurement data and estimate a single value. They can be coupled with Tikhonov regularization techniques that consider prior information, but they do not model uncertainty on the measurement data. Due to their averaging nature, such methods will typically work well if measurement errors are comparable (i.e., there are no outliers) and do not suffer from a constant bias (i.e., they are evenly spread around zero). They may work poorly in case of such a systematic bias, and will certainly be very sensitive to the presence of outliers.

1.3 Non-deterministic inverse problem based on the probability theory

In this section, we discuss the inverse method based on the probability theory. Probability theory [Jaynes, 2003, Zio and Pedroni, 2013] is one of the tools traditionally used to represent uncertainty in risk assessment. Probability is a measure of the occurrence of an event, and it is a single-valued measure of uncertainty, i.e., a single number $p(A)$ represents the uncertainty about the occurrence of an event A. Different interpretations of probabilities exist, such as the relative frequency (the frequentist view) and the subjective or Bayesian (the subjective (Bayesian) view). For the first one, the probability is the fraction of times an event A occurs if the situation is to occur an infinite number of times. For the latter, the probability of an event A represents the degree of belief of the assigner for the occurrence of A. Next, we will discuss the probability theory in uncertainty representation and propagation.

1.3.1 Uncertainty representation in probability theory

In the probability theory, a random variable represents the uncertainty of the random events. Consider Ω to be the sample space of a random events, which is a set of all possible outcomes. A real random variable X is a measurable mapping:

$$X : \Omega \longrightarrow E \subseteq \mathbb{R}$$

The random variable is discrete when E is discrete; otherwise, it is continuous. A real random vector $\boldsymbol{X}$ is a measurable mapping:

$$\boldsymbol{X} : \Omega \longrightarrow \boldsymbol{E} \subseteq \mathbb{R}^d$$

The real random vector is a vector whose components are random variables. In probability theory, random variable X is used to model uncertainties. These can be simple real-valued random variables or more complex objects such as stochastic processes or random fields [Batou and Soize, 2013, Gong et al., 2020].

Now we will discuss the uncertainty propagation in probability theory in the case of the forward problem. We consider a model of the form:

$$\boldsymbol{Y} = f(\boldsymbol{X})$$

where $\boldsymbol{X} = (X_1, ..., X_d)$ is a random vector with a given probability law, and f is a function which associates a value x of $\boldsymbol{X}$ to a value y of the variable of interest $\boldsymbol{Y}$. The choice of a probability law or distribution for $\boldsymbol{X}$ depends on the availability of data: 1) If no data are available, we can introduce some a priori information to directly construct a probability law for $\boldsymbol{X}$, e.g., with the maximum entropy principle [Harremoës and Topsoe, 2001, De Martino and De Martino, 2018]. 2) If data are available, we can use statistical inference methods. Monte Carlo simulation (MCS) is one of the techniques [Anderson, 1976, Couto et al., 2013, Cunha et al., 2014] used for the propagation of the uncertainties from X to Y. Next, we will discuss the inverse method based on the probability theory.

1.3.2 Inverse method based on Bayesian Inference

In this section, we discuss the non-deterministic inverse method based on Bayesian inference. In this method, uncertainty on the information is modeled with the probabilistic approach. In the context of Bayesian inference, the main idea is to find the posterior probability distribution of model parameters $\boldsymbol{\theta}$ given the observed measurement data $\tilde{\boldsymbol{u}}$. The result is defined by applying Bayes' formula:

$$p(\boldsymbol{\theta} \mid \tilde{\boldsymbol{u}}) = \frac{p(\tilde{\boldsymbol{u}} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})}{\int p(\tilde{\boldsymbol{u}} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{1.6}$$

where $p(\tilde{\boldsymbol{u}} \mid \boldsymbol{\theta})$ is the likelihood of measurement data, i.e., $p(\tilde{\boldsymbol{u}} \mid \boldsymbol{\theta})$ depends on the probability distribution of the error $\varepsilon = \boldsymbol{u} - \tilde{\boldsymbol{u}}$ on the measurements, $p(\boldsymbol{\theta})$ is the prior probability distribution of model parameters which corresponds to the prior knowledge on $\boldsymbol{\theta}$. As the denominator term in Equation ((1.6)) is a constant, i.e.,

$c = p(\tilde{\boldsymbol{u}}) = \int p(\tilde{\boldsymbol{u}} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$, a simplified form of the above formulation is

$$p(\boldsymbol{\theta} \mid \tilde{\boldsymbol{u}}) \propto p(\tilde{\boldsymbol{u}} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \tag{1.7}$$

The idea of identification in the Bayesian inverse method is to generate a posterior distribution of parameters from the probability distribution of measurements and prior knowledge about parameters. To estimate a posterior distribution, one of the ways is the Markov chain Monte Carlo(MCMC) sampling method [van Ravenzwaaij et al., 2016]. Bayesian inverse methods find their application in various fields such as mechanical parameter identification [Emery et al., 2016], finance [Ota et al., 2019], and artificial intelligence [Tipping, 2004]. In the context of identification of mechanical models, Bayesian inverse methods have been used, for example, for the identification of elastic properties from full-field displacement measurements [Gogu et al., 2010]. Unlike the deterministic least-squares approach, the Bayesian inverse method results in a probability distribution to represent the possible values of the identified parameter. The key feature of the Bayesian inverse method is that it takes prior information about parameters using a probability distribution. A way to choose the prior distribution can be from subjective engineering judgment or maximum entropy principle. The choice of the prior distribution plays a vital role in Bayesian inverse method. However, in the case of incomplete knowledge or partial ignorance, researchers have debated a lot about the limitations of the probability theory to represent uncertainty [Moens and Hanss, 2011, Ferson and Ginzburg, 1996, Helton and Johnson, 2011]. [Hose and Hanss, 2019] describes the limitation of probability theory to represent partial ignorance with the following example: "Suppose, you have the same knowledge about a certain positive variable $x > 0$, e.g. the proportion of alcohol in a liquid, and about its reciprocal $y = \frac{1}{x}$. Then, $x \in [a, b]$ is equivalent to $y \in [\frac{1}{b}, \frac{1}{a}]$, but a uniform distribution on $[a, b]$ is not compatible with a uniform distribution on $[\frac{1}{b}, \frac{1}{a}]$. Therefore, uniform probability distributions cannot represent total ignorance since invariance under transformations cannot be guaranteed".

Even though probability theory is well suited to model aleatory uncertainty but in the case of incomplete information related to measurements or parameters, Bayesian inverse methods can give false predictions [Chen et al., 2000]. Bayesian methods can be very efficient and accurate if the model is right. However, as for the least-square methods, Bayesian methods can be quite sensitive to outliers (as we shall see in the experiments of Chapter 3) and require a strong modeling effort to be accurate and not be subject to model misspecification biases. In this sense, their robustness can be limited, even if they are the first choice when it comes to

model and process measurement uncertainty.

1.4 Non-deterministic inverse problem based on the non-probabilistic approach

In the last section, we discussed the limitation of probability theory to represent imprecise or incomplete information. Alternative non-probabilistic approaches are proposed in the literature to represent uncertainty on the data irrespective of its nature or source, and it includes imprecise probabilities [Walley, 1991], intervals [Moore, 1966], fuzzy sets [Zadeh, 1965], possibility theory [Baudrit and Dubois, 2006] and theory of belief function (evidence theory [Shafer, 1976]). They play an essential role in solving the inverse and the forward problem taking into account uncertainty propagation. Concerning the non-probabilistic approaches, we will limit ourselves to the intervals and sets to represent uncertainty as our work focuses on this particular case. However, intervals are the basis of the other theories and extensions to these theories is left for future investigations.

1.4.1 Intervals to represent uncertainty

Within the framework of interval analysis, an interval $[x]$ in $\mathbb{R}$ is a closed set of connected real values noted by $[x] = [\underline{x}, \overline{x}] = \{x \in \mathbb{R} \mid \underline{x} \leq x \leq \overline{x}\}$ where $\underline{x} \in \mathbb{R}$ is the lower bound and $\overline{x}$ is the upper bound [Jaulin et al., 2001]. In the context of uncertainty representation, unlike probability theory, with intervals the idea is that the actual value of the uncertain quantity is between two bounds without requiring information on the likelihood of each value within that interval. In our work, we choose to describe uncertainty on the measurements in interval form, as such a description requires almost no assumption regarding the nature and source of uncertainty [Zio and Pedroni, 2013]. To describe prior information about parameters, we use a multidimensional extension of intervals, i.e. hypercube or box of $\mathbb{R}^n$ defined as the Cartesian product of n intervals. The set of all n-dimensional boxes will be denoted by $\mathbb{IR}^n$. For example, in the case of two parameters, x_1 and x_2, information on them is described by set $\mathbb{X}$ such that $\mathbb{X} = [\boldsymbol{x}] = [x_1] \times [x_2] = [\underline{x_1}, \overline{x_1}] \times [\underline{x_2}, \overline{x_2}]$ as illustrated in Figure 1.1. Boxes correspond to the multidimensional sets that are easiest to describe sets.

How intervals can be used to solve the forward or inverse problem apart from their role as the uncertainty representation? The answer to this question is uncertainty propagation with intervals. We discuss this in the next section.

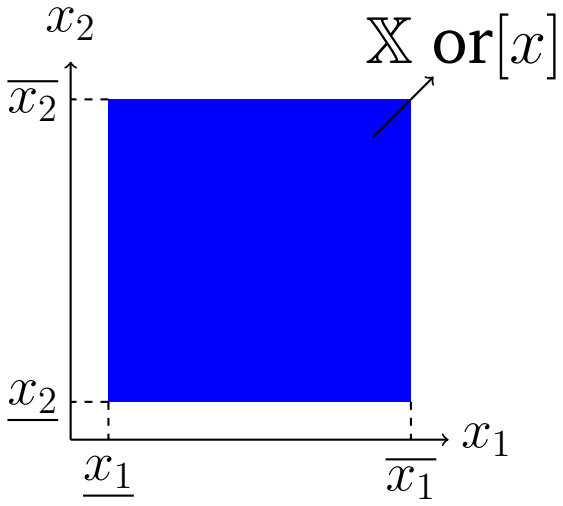


Figure 1.1 – A box $[\boldsymbol{x}]$ of $\mathbb{IR}^2$

1.4.2 Uncertainty propagation with intervals

We consider a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ as

$$y = f(x) \tag{1.8}$$

The function f extends to its interval form as

$$[\boldsymbol{y}] = [\boldsymbol{f}]([\boldsymbol{x}]) \tag{1.9}$$

where $[f]$ is an interval function. The role of this function in the context of the forward interval uncertainty propagation is to obtain quantity y in interval form with input quantity x in the interval form. The commonly used method for interval propagation includes interval arithmetic [Moore, 1966], optimization [Moens and Hanss, 2011], the vertex method [Dong and Shah, 1987, Akpan et al., 2001], perturbation method [Chen et al., 2002, Jaulin et al., 2001], Interval finite element method [Muhanna and Mullen, 2001, Zhang, 2005]. Before discussing them, we try to give some notions related to intervals.

The width of any non-empty interval $[x]$ is defined as

$$w([x]) = \overline{x} - \underline{x} \tag{1.10}$$

The midpoint or centre of any bounded and non-empty interval $[x]$ is defined as

$$m([x]) = \frac{\underline{x} + \overline{x}}{2} \tag{1.11}$$

In the context of set-theoretic operations, the intersection of two intervals $[x_1]$ and $[x_2]$ is given as

$$[x_1] \cap [x_2] = \{z \in \mathbb{R} \mid z \in [x_1] \text{ and } z \in [x_2]\} = [max\{\underline{x_1},\ \underline{x_2}\}, min\{\overline{x_1},\ \overline{x_2}\}] \quad (1.12)$$

The union of two intervals $[x_1]$ and $[x_2]$ is given as

$$[x_1] \cup [x_2] = \{z \in \mathbb{R} \mid z \in [x_1] \text{ or } z \in [x_2]\} = [min\{\underline{x_1},\ \underline{x_2}\}, max\{\overline{x_1},\ \overline{x_2}\}] \quad (1.13)$$

1.4.2.1 Interval arithmetic

The fundamental arithmetic operations [Moore, 1966] for non-empty closed intervals can be defined as:

$$[x_1] + [x_2] = [\underline{x_1} + \underline{x_2},\ \overline{x_1} + \overline{x_2}] \quad (1.14)$$

$$[x_1] - [x_2] = [\underline{x_1} - \overline{x_2},\ \overline{x_1} - \underline{x_2}] \quad (1.15)$$

$$[x_1] * [x_2] = [min\{\underline{x_1 x_2}, \underline{x_1}\overline{x_2}, \overline{x_1}\underline{x_2}, \overline{x_1 x_2}\},\ max\{\underline{x_1 x_2}, \underline{x_1}\overline{x_2}, \overline{x_1}\underline{x_2}, \overline{x_1 x_2}\}] \quad (1.16)$$

$$\begin{aligned} 1/[x_2] &= \emptyset \ \text{ if } [x_2] = [0,0], \\ 1/[x_2] &= [1/\overline{x_2},\ 1/\underline{x_2}] \ \text{ if } 0 \notin [x_2], \end{aligned} \quad (1.17)$$

$$[x_1]/[x_2] = [x_1] * 1/[x_2] \quad (1.18)$$

In many problems, these simple interval arithmetic operations can be used in a straightforward way to propagate intervals through functions to obtain the interval output. Interval arithmetic is a very convenient tool, but it is not without drawback, in particular in terms of the conservatism of the obtained result [Zhang, 2005, Jaulin et al., 2001]. Consider a function $f(x) = x * x + x$ from $\mathbb{R}$ to $\mathbb{R}$, and the aim is to propagate interval $[x] = [-1, 1]$ through it. If we use interval arithmetic technique, the output of this function is $[-2, 2]$. We can express the same function in its equivalent form $(x + \frac{1}{2})^2 - \frac{1}{4}$, and if we propagate interval $[x] = [-1, 1]$ through it, then the output of this function is $[-\frac{1}{4}, 2]$. If we compare the two outputs, then the width of the second one is smaller than the first one, i.e., the output result is less conservative. The reason for getting an overestimated result is that there are multiple occurrences of the interval variable $[x]$ in the first form of the function. Since interval arithmetic treats each interval variable as different and independent, it can not take into account dependency of the same interval variable when it occurs several times in the expression. Many authors [Sofi and Romeo, 2016, Zhang, 2005, Moore, 1966, Faes and Moens, 2019] called this dependency phenomenon. Nevertheless, this interval arithmetic techniques can be useful in

some cases.

1.4.2.2 Optimization method

As discussed in the previous section, interval arithmetic has some limitations to propagate interval through the model in some cases. Another way to solve the interval problem without using interval arithmetic is the use of optimization method [Moens and Hanss, 2011]. Consider a function $y = f(x)$ from $\mathbb{R}$ to $\mathbb{R}$, and we want to propagate interval $[x]$ through it to obtain interval output of y, i.e., the range of $[y]$. In this method, to calculate range, $[y]$, the function is optimized at possible values of x that falls within the range of $[x]$. The range of $[y]$ as an optimization problem is given as

$$\begin{aligned} \underline{y} &= \min_{x \in [x]} f(x) \\ \overline{y} &= \max_{x \in [x]} f(x) \end{aligned} \tag{1.19}$$

In this method, we avoid the dependence problem easily as we evaluate the function deterministically with no interval propagation explicitly. This method can be useful when the function under evaluation has a non-monotonic nature. However, with practical engineering problem such as industrial Finite Element problem, the number of function evaluations increases with the number of input interval parameters which then leads to high computational cost for the optimization problem.

1.4.2.3 Vertex method

Another way to solve the interval problem without using interval arithmetic is the use of the vertex method [Dong and Shah, 1987, Akpan et al., 2001]. If we want to propagate interval $[x]$ through a function $y = f(x)$ from $\mathbb{R}$ to $\mathbb{R}$ to obtain the range of $[y]$ then in this method, the function is evaluated at possible values of x that falls within the range of $[x]$. The range of $[y]$ with the vertex method is given as

$$[\underline{y}, \overline{y}] = [min(f(\underline{x}), f(\overline{x})), max(f(\underline{x}), f(\overline{x}))] \tag{1.20}$$

In this method, we avoid the dependence problem easily as the interval propagation does not happen explicitly. This method can be useful only when the function under evaluation has a monotonic nature, otherwise, we only obtain an inner approximation of the true result. However, the number of function evaluations increases with the number of input interval parameters that leads to high computational cost.

1.4.2.4 Perturbation method

Another way to solve the interval problem without using interval arithmetic is the use of a perturbation method [Chen et al., 2002, Jaulin et al., 2001], which aims at obtaining a tight enclosure of the interval output. In this method, we approximate the function $f(x)$ from $\mathbb{R}$ to $\mathbb{R}$ evaluation around the middle point of the input interval parameter, i.e., $m[x]$ with the first-order Taylor series expansion

$$f(x) \approx f(m[x]) + f'(x)(x - m[x]) \tag{1.21}$$

This approximated function can then be used with the optimization problem(1.19) to know the interval output.The perturbation method with a first-order approximation is only applicable when the width of the input interval parameter is small enough to obtain the bounded output. Higher-order approximation helps to improve the performance of the perturbation method; however, this may entail cumbersome calculations.

1.4.2.5 Interval Finite Element method

The Finite Element(FE) method is a widely used method for solving problems of engineering such as structural mechanics, heat transfer, fluid flow, mass transport, and electromagnetic potential. It is a mathematical tool to solve partial differential equations representing the physics. Let us consider a simple example of linear elasticity of 1D bar with 2 elements and 3 nodes. The FE formulation [Fish and Belytschko, 2007] of this problem results in a linear system of equations as:

$$\boldsymbol{KU} = \boldsymbol{F} \tag{1.22}$$

where $\boldsymbol{K}$ is the global stiffness matrix, $\boldsymbol{U}$ is the vector of nodal displacements, and $\boldsymbol{F}$ is a force vector. Its interval linear system is given as:

$$[\boldsymbol{K}][\boldsymbol{U}] = [\boldsymbol{F}] \tag{1.23}$$

The global stiffness matrix $[\boldsymbol{K}]$ is obtained from assembling two elemental stiffness matrices for this particular example as:

$$[\boldsymbol{K}] = [\boldsymbol{K_1}] + [\boldsymbol{K_2}] = \begin{bmatrix} [\underline{K_{11}}, \overline{K_{11}}] & [\underline{K_{12}}, \overline{K_{12}}] & 0 \\ [\underline{K_{21}}, \overline{K_{21}}] & [\underline{K_{22}}, \overline{K_{22}}] + [\underline{K_{22}}, \overline{K_{22}}] & [\underline{K_{23}}, \overline{K_{23}}] \\ 0 & [\underline{K_{32}}, \overline{K_{32}}] & [\underline{K_{33}}, \overline{K_{33}}] \end{bmatrix} \tag{1.24}$$

where $\boldsymbol{K_1}$ and $\boldsymbol{K_2}$ are given as

$$[\boldsymbol{K_1}] = \begin{bmatrix} [\underline{K_{11}}, \overline{K_{11}}] & [\underline{K_{12}}, \overline{K_{12}}] & 0 \\ [\underline{K_{21}}, \overline{K_{21}}] & [\underline{K_{22}}, \overline{K_{22}}] & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$[\boldsymbol{K_2}] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & [\underline{K_{22}}, \overline{K_{22}}] & [\underline{K_{23}}, \overline{K_{23}}] \\ 0 & [\underline{K_{32}}, \overline{K_{32}}] & [\underline{K_{33}}, \overline{K_{33}}] \end{bmatrix}$$

When we represent uncertainty on stiffness parameter with intervals, during the assembling process in Equation (1.24), we can observe stiffness parameter $[K_{22}]$ occurred twice. In the FE problems, as the number of nodes and elements increases, the size of the global stiffness matrix increases that can lead to multiple occurrences of the stiffness parameters. As discussed in section 1.4.2.1, propagation of the interval stiffness parameter through Equation (1.22) by interval arithmetic will provide overly conservative results, due to dependency issues. To tackle the issue of dependency, in the context of Interval FE method, [Muhanna and Mullen, 2001, Zhang, 2005] have proposed Element-by-Element(EBE) method, which is useful in assembling the global stiffness matrix. In this method, the idea is to detach elements so that there are no connections between elements to avoid element coupling in the element assembly procedure. This method has shown to provide sharp bounds on interval output in the context of the Interval FE method. The Interval FE method consists of solving a linear interval system of equations. [Neumaier, 1990, Shary, 2001, Garloff, 2008] have proposed methods such as interval Gauss-seidel method, or interval Gaussian elimination to solve the linear interval system.

1.4.3 Non-deterministic inverse problem based on the interval theory

In this section, we review existing inverse methods based on intervals and sets.

1.4.3.1 Set inversion based approach to solve the inverse problem

Consider a function f from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $\mathbb{Y}$ a subset of $\mathbb{R}^m$, then the set inversion [Jaulin et al., 2001] problem is defined as the characterization of the set

$$\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \in \mathbb{Y}\} = f^{-1}(\mathbb{Y}) \tag{1.25}$$

For function f, the set $\mathbb{Y}$ is represented as $[\boldsymbol{y}]$, which is a known interval vector of m measurements, and the set $\mathbb{X}$ which is to be characterized is represented as $[\boldsymbol{x}]$, which is a box to which $\mathbb{X}$ is guaranteed to belong. A way to approximate the compact set $\mathbb{X}$ in a guaranteed way is the use of subpavings. A subpaving of a box $[\boldsymbol{x}]$ is the union of non-empty and non-overlapping sub boxes of $[\boldsymbol{x}]$. The inverse set solution approximates $\mathbb{X}$ between two subpavings $\underline{\mathbb{X}}$ and $\overline{\mathbb{X}}$ such that $\underline{\mathbb{X}} \subseteq \mathbb{X} \subseteq \overline{\mathbb{X}}$ using SIVIA (Set inversion via Interval Analysis) method [Jaulin and Walter, 1993, Braems et al., 2001] which is described in Algorithm 1.

Algorithm 1: SIVIA

Input: f, $[\boldsymbol{y}]$, $[\boldsymbol{x}](0)$, ε

Output: $\mathbb{X}$ such that$\underline{\mathbb{X}} \subseteq \mathbb{X} \subseteq \overline{\mathbb{X}}$

1 **begin**
2 $\quad \underline{\mathbb{X}} := \emptyset$; // Initialization
3 $\quad \overline{\mathbb{X}} := \emptyset$; // Initialization
4 $\quad \mathscr{L} := \{[\boldsymbol{x}](0)\}$;
5 $\quad$ Pull $[\boldsymbol{x}]$ from $\mathscr{L}$;
6 $\quad$ **if** $[f]([\boldsymbol{x}]) \cap [\boldsymbol{y}] = \emptyset$ **then**
7 $\quad\quad$ return; // Figure 1.2(b)
8 $\quad$ **if** $[f]([\boldsymbol{x}]) \subset [\boldsymbol{y}]$ **then**
9 $\quad\quad$ $\underline{\mathbb{X}} := \underline{\mathbb{X}} \cup [\boldsymbol{x}]$; $\overline{\mathbb{X}} := \overline{\mathbb{X}} \cup [\boldsymbol{x}]$; return; // Figure 1.2(c)
10 $\quad$ **if** $w([\boldsymbol{x}]) < \varepsilon$ **then**
11 $\quad\quad$ $\overline{\mathbb{X}} := \overline{\mathbb{X}} \cup [\boldsymbol{x}]$; return; // Figure 1.2(d)
12 $\quad$ **else**
13 $\quad\quad$ bisect $[\boldsymbol{x}]$ and push into $\mathscr{L}$;
14 $\quad$ **if** $\mathscr{L} \neq \emptyset$ **then**
15 $\quad\quad$ go to line 5
16 $\quad$ **return** $\mathbb{X}$;

SIVIA method first requires a search box $[\boldsymbol{x}](0)$ of parameters to which $\overline{\mathbb{X}}$ belongs. The search process then needs to create the sub boxes of $[\boldsymbol{x}](0)$. The sub boxes can be created with regular subpaving [Tornil-Sin et al., 2010, Jaulin et al., 2001] of $[\boldsymbol{x}](0)$, i.e., each of its boxes obtained from a finite succession of bisections and selections. Figure 1.2 describes the basic cases of SIVIA using Algorithm 1. They are explained as follows.

1. If $[f]([\boldsymbol{x}]) \supset [\boldsymbol{y}]$, then $[\boldsymbol{x}]$ may contain a part of the solution set (see Figure 1.2(a)). In this situation $[\boldsymbol{x}]$ is said to be undertermined. If $w([\boldsymbol{x}])$ is greater than a precision parameter ε, then it should be bisected (line 10 and line 13) so that we again continue to search the solution set of parameters

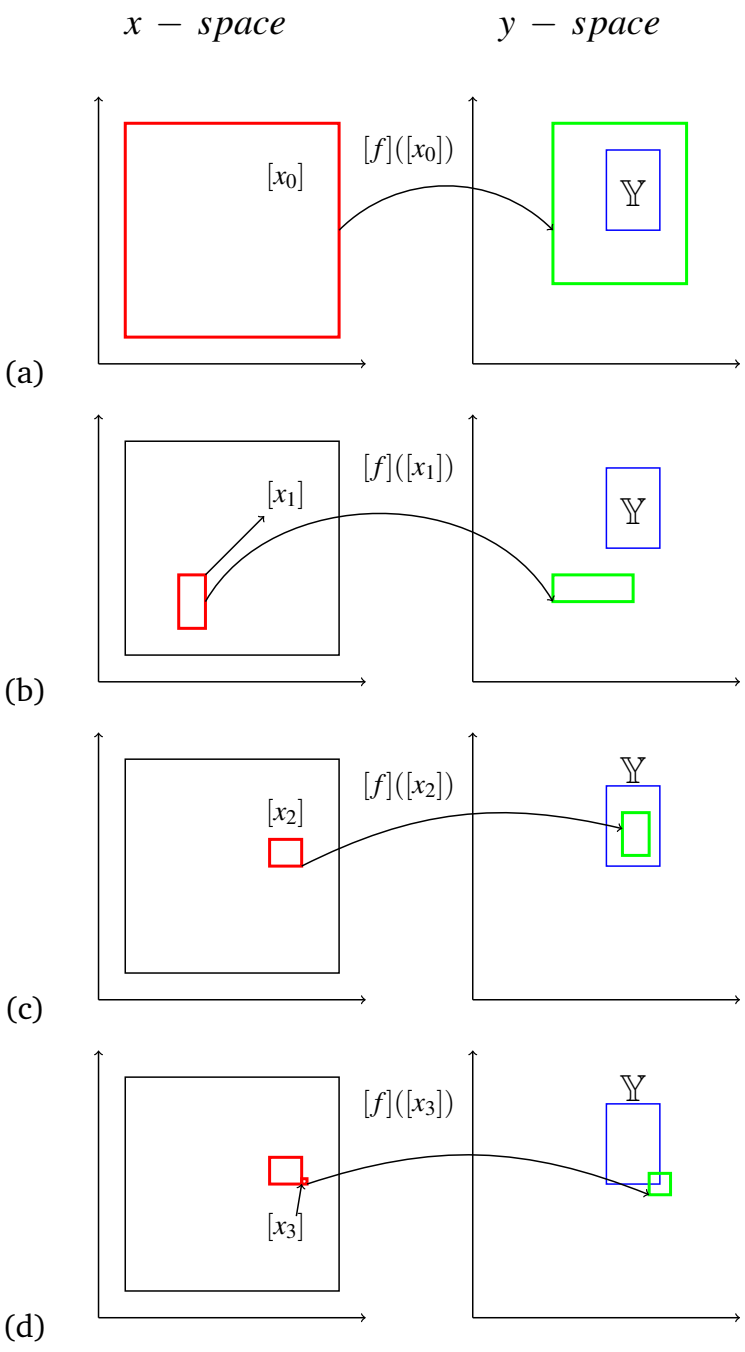


Figure 1.2 – SIVIA cases (a) the box $[\boldsymbol{x_0}]$ to be checked is undetermined and will be bisected; (b) the box $[f]([\boldsymbol{x_1}])$ does not intersect $\mathbb{Y}$ and $[\boldsymbol{x_1}]$ is rejected; (c) the box $[f]([\boldsymbol{x_2}])$ is included in $\mathbb{Y}$; and $[\boldsymbol{x_2}]$ is accepted; (d) the box $[\boldsymbol{x_3}]$ is undetermined but to small to be bisected but stored in $\overline{\mathbb{X}}$

from line 5.

2. If $[f]([\boldsymbol{x}]) \cap [\boldsymbol{y}] = \emptyset$(see Figure 1.2(b)), then $[\boldsymbol{x}]$ does not belong to $\mathbb{X}$ and is excluded from further bisection.

3. If $[f]([\boldsymbol{x}]) \subset [\boldsymbol{y}]$(see Figure 1.2(c)), then $[\boldsymbol{x}]$ does belong to solution subpaving $\mathbb{X}$, and is stored in $\underline{\mathbb{X}}$ and $\overline{\mathbb{X}}$.

4. The last case is described in Figure 1.2(d). If the box considered is determined, but its width is lower than ε, then it is stored in the outer approximation $\overline{\mathbb{X}}$ of $\mathbb{X}$.

SIVIA method finds its application in robotics and control system [Jaulin et al., 2001] to identify guaranteed enclosure of parameters when uncertainty on measurements is described with intervals. To compute $[f]([\boldsymbol{x}])$, the SIVIA method makes use of interval uncertainty propagation methods described in the section 1.4.2 and they also affect the final quality(overestimation) of the solution set. The complexity of the SIVIA method quickly increases as the number of parameters to identify increases, which is why it is not really well adapted to high-dimensional mechanical problems

The presented SIVIA method gives an empty solution set of parameters in the presence of outliers in the measurements. The fact that $[f]([\boldsymbol{x}]) \cap [\boldsymbol{y}] = \emptyset$ can be due to the fact that at least one measurement is inconsistent from the box $[\boldsymbol{y}]$ is well-known in interval analysis. A common method to solve such issues is to consider the so-called q-intersection method [Drevelle and Bonnifait, 2012, Sandretto et al., 2014]. Given a set $\mathcal{X} = \{\mathbb{X}^1, \ldots, \mathbb{X}^m\}$ of set-valued solutions corresponding to $[y_1]$ to $[y_m]$ interval measurements, their q-intersection consists in applying the formula

$$\overset{q}{\mathbb{X}} = \bigcup_{E \subseteq \mathcal{X}, |E|=q} \bigcap_{\mathbb{X}_i \in E} \mathbb{X}_i, \tag{1.26}$$

that is to take the union over all possible intersections of q elements. We illustrate the q-relaxed intersection method (see Figure 1.3). Consider a set $\mathcal{X} = \{\mathbb{X}^1, \ldots, \mathbb{X}^4\}$ of set-valued solutions corresponding to $[y_1]$ to $[y_4]$ interval measurements. If all the four measurements are consistent with one another we get the non-empty solution set (see Figure 1.3(a)). If all the four measurements are not consistent with one another then we get the empty solution set (see Figure 1.3(b)). To obtain non-empty solution set, if we use q-relaxed intersection method with q=2 then we get final result with $[y_2]$ considered as outlier. While such an approach brings good results if the number m is not too large, and if q is close to m, it

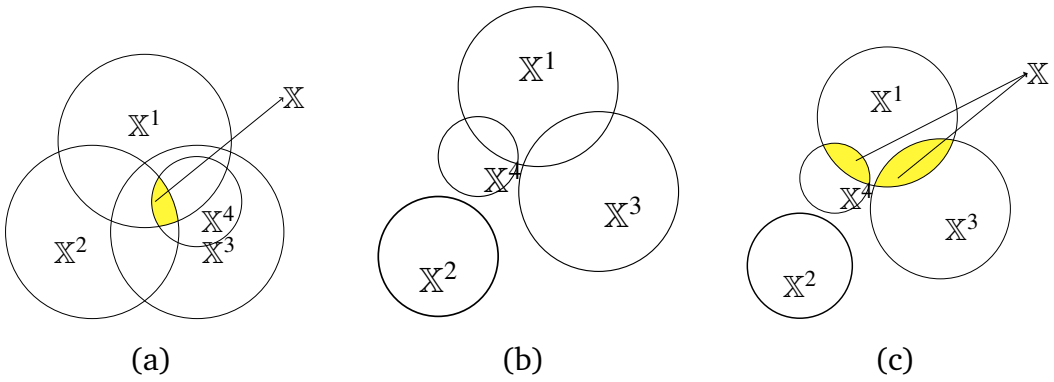


Figure 1.3 – q-relaxed intersection method (a) all the four measurements are consistent with one another; (b) all the four measurements are not consistent with one another; (c) non-empty solution set with q=2

becomes quickly intractable when the number m of sets to merge increases, as the complexity of applying Equation (2.6) increases factorially with the number $m - q$. Therefore, in the setting of inverse mechanical problems where the minimal number of measurements is some dozens and where it is not unusual to have thousands of measurements (for large mechanical structures), the q-intersection method is not applicable.

1.4.3.2 Inverse problem based on the Interval Finite Element method

To solve the mechanical structural inverse problem under the presence of uncertain interval measurements, [Xiao, 2005, Fedele et al., 2012, Xiao et al., 2013] have proposed methods based on the Interval Finite Element(FE) method. For a linear static structure, the linear system of equations after Finite Element formulation is given as:

$$\boldsymbol{KU} = \boldsymbol{F} \tag{1.27}$$

The corresponding Interval FE linear system is given as

$$[\boldsymbol{K}][\boldsymbol{U}] = [\boldsymbol{F}] \tag{1.28}$$

The inverse Interval FEM problem consists of estimating the tight enclosure of stiffness parameters from interval displacement measurement. Like forward Interval FE method, inverse Interval FE method also faces the issue of the overestimation in the output and dependency phenomenon (assembly of global stiffness matrix). It is solved in the same way as in the case of the former. The Interval FE method is still a work in progress [Sofi and Romeo, 2016, Faes and Moens, 2019] in the context of the forward or interval uncertainty propagation. Recently, in the context

of Interval FEM, [Faes and Moens, 2019] have proposed an interval field approach to take into account the dependency of the interval variables.

1.4.3.3 Discussion about SIVIA and inverse Interval FE method

Both SIVIA and inverse Interval FE methods are inverse methods based on the non-probabilistic representation of uncertainty. They make use of interval uncertainty propagation techniques in solving the inverse problem and focus on obtaining tight enclosure of the parameters to identify the solution set. SIVIA method has the advantage of providing guarantees, in the sense that the approximation is known to contain the actual solution set, but can be computationally intensive, especially for high-dimensional problems. The inverse Interval FE method is specifically introduced to solve the mechanical inverse problem. Still, it has some limitations to solve the problem of identification with a large number of measurements because of its complex Interval FEM formulation [Xiao, 2005, Sofi and Romeo, 2016]. In practice, the mechanical inverse problem consists of solving the problem of identification with a large number of measurements, e.g., full-field displacement measurements. Both methods are useful in situations where one is interested in identifying parameters with a tight enclosure from a limited number of measurements. Still, they have limitations in terms of identification with inconsistent interval measurements when they are in a large amount.

1.5 Summary

In this chapter, we presented some of the most common and recent deterministic and non-deterministic inverse methods. For the deterministic way of solving the inverse problem, we discussed its limitations in solving the inverse problem in the presence of highly inconsistent measurement data. Next, we discussed the non-deterministic Bayesian inverse method, a probabilistic approach. It has an advantage when sufficient data are available to solve an inverse problem. Still, in the case of incomplete information, they have certain limitations, such as susceptibility to the presence of outliers in the data, they often require to make extra assumptions of Bayesian model that may be false, hence may lead to misleading results. Our thesis work focuses on the development of the inverse strategy in the case of non-probabilistic uncertainty representation method. We presented interval theory as a means to model uncertainty on the information, and interval propagation methods in the forward and inverse problem. We reviewed existing inverse strategy based on interval theory such as SIVIA and Interval FEM and their limitations with dealing

with a large number of inconsistent measurements. In our thesis, we continue our research on an inverse strategy based on the interval approach to deal with inconsistent measurements in a large number, and strategy presented in the next chapter.

Inverse problem strategy based on interval uncertainty representation approach.

Contents

2.1 Introduction

In this chapter, we explore how interval-based approaches can help to obtain a solution to the inverse problem, mainly when measurements are inconsistent with one another. We use intervals that allow us to take into account both prior information about the parameters as well as measurement uncertainty in the form of sets (interval or boxes) during the inversion process and we construct a solution set.

2.2 Preliminaries and related works

We introduce the set-based solution to the inverse problem in Section 2.2.1, and then detail how the problem of inconsistent measurements is commonly solved in the literature in Section 2.2.2.

2.2.1 Set-valued inverse problem

We consider an inverse problem where we want to identify some parameters of a model $y = f(\theta)$ from measurements made on quantity $y \in \mathbb{R}^N$. The model f yields the relationship between the M model parameters $\theta \in \mathbb{R}^M$ and the measured quantity, under given experimental conditions. We will denote by $\tilde{y} \in \mathbb{R}^N$ the measurements made on y. In this section, we consider that the model is accurate, in the sense that any discrepancy between $f(\theta^*)$, θ^* being the true parameter values, and $\tilde{y}$ is due to some measurement errors, i.e., noise, systematic bias, etc. This means that we leave the issue of integrating model error for future investigations.

In this work, we will consider that our uncertainty on any real-valued measurement $\tilde{y}$ is provided by an interval $[y]$ [Jaulin et al., 2001], that is a closed set of connected real values noted by $[y] = [\underline{y}, \overline{y}] = \{y \in \mathbb{R} \mid \underline{y} \leq y \leq \overline{y}\}$ where $\underline{y}$ and $\overline{y}$ are respectively the lower and the upper bounds of the interval. One advantage of using intervals is that it requires almost no assumption regarding the nature and source of uncertainty [Zio and Pedroni, 2013].

In particular, our uncertainty on each measurement $\tilde{y}_k$ will be described by such an interval, and the overall uncertainty on all measurements will correspond to a hyper-cube $\mathbb{S}_y$

$$\mathbb{S}_y = \prod_{k=1}^{N} [\underline{\tilde{y}_k}, \overline{\tilde{y}_k}] \subset \mathbb{R}^N \tag{2.1}$$

where each measurement is described by its lower bound $\underline{\tilde{y}_k}$ and upper bound $\overline{\tilde{y}_k}$.

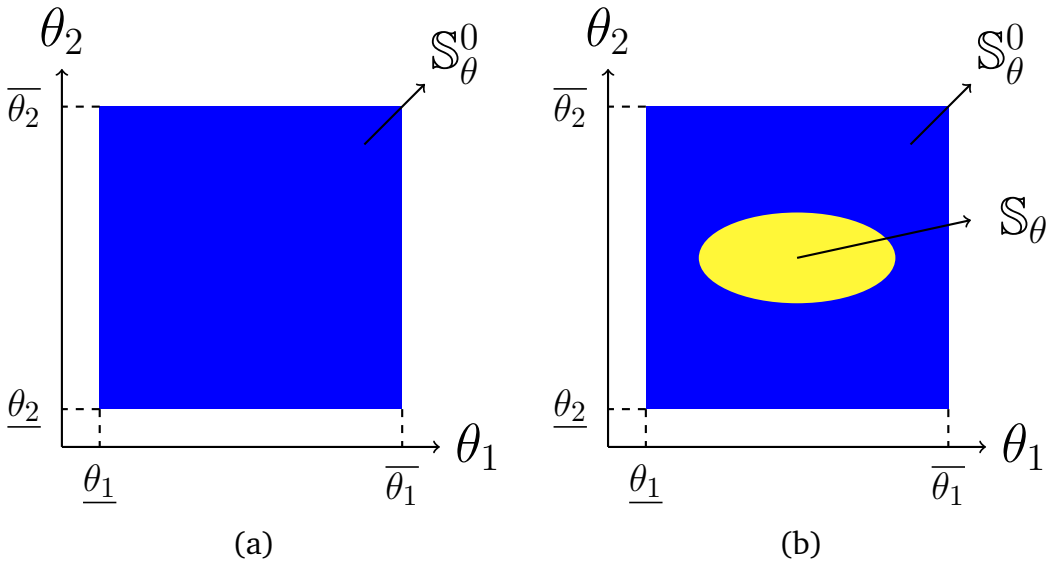


Figure 2.1 – Set description through a box. (a) Prior set ($\mathbb{S}_\theta^0$) (b) Solution set ($\mathbb{S}_\theta$).

Similarly, we will assume that our prior information about the parameters θ is provided by a hyper-cube $\mathbb{S}_\theta^0 \subset \mathbb{R}^M$ that gives a simple description of the physical boundaries in which θ can lie (see Figure 2.1(a)). The set-valued solution $\mathbb{S}_\theta$ to the inverse problem can then be simply described (see Figure 2.1(b)) as the set of all parameter values within $\mathbb{S}_\theta^0$ that are consistent with the observed uncertain measurements, i.e.,

$$\mathbb{S}_\theta = \{\theta \in \mathbb{S}_\theta^0 \mid f(\theta) \in \mathbb{S}_y\}. \tag{2.2}$$

Since in the current approach all measurements are considered independent of each other, due to the fact that we take their Cartesian product, computing $\mathbb{S}_\theta$ can alternatively be written as the result of the intersection (see Figure 2.2)

$$\mathbb{S}_\theta = \bigcap_{k=1}^{N} \mathbb{S}_\theta^k \tag{2.3}$$

where

$$\mathbb{S}_\theta^k = \{\theta \in \mathbb{S}_\theta^0 \mid f_k(\theta) \in [\underline{\tilde{y}_k}, \overline{\tilde{y}_k}]\} \text{ s.t } f_k(\theta) \in \mathbb{R} \tag{2.4}$$

is the set of parameter values consistent with the k^{th} measurement. In general, obtaining $\mathbb{S}_\theta^k$ can be done through some set-inversion algorithm, for instance the SIVIA algorithm [Jaulin et al., 2001], or grid-based approaches as described in Section 2.4.

Example 1. We illustrate the set-valued inverse problem on a toy example. We consider a spring-mass system shown in Figure 2.3 which can be described by the

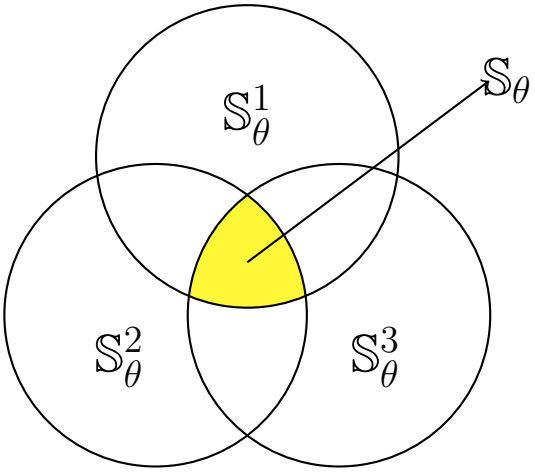


Figure 2.2 – Intersection of solution sets

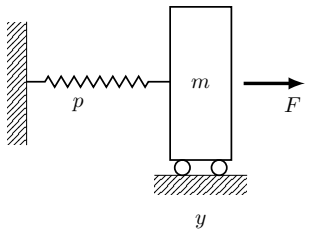


Figure 2.3 – Spring-mass system

equation

$$F/p = f(\theta) = y \tag{2.5}$$

where F represents the force applied on the spring in $Newton(N)$, p is the spring stiffness constant (N/m), is the parameter to estimate. y is the measured displacement of the spring in meter (m). We consider a case where a force F=100 N is applied on the spring and a displacement $\tilde{y}_1$=0.01 m is measured. Here, the inverse problem consists of determining the parameter from the measurement $\tilde{y}_1$. To do so, we describe uncertainty on the prior knowledge about the parameter ($\theta = p$) and the measurement in the interval form such that

$$\mathbb{S}_\theta^0 = [P] = [8000\ , 12000]$$

and our uncertainty $\tilde{y}$ on the measurement as

$$\mathbb{S}_y = [\underline{\tilde{y}_k}, \overline{\tilde{y}_k}] = [0.009\ , 0.0110].$$

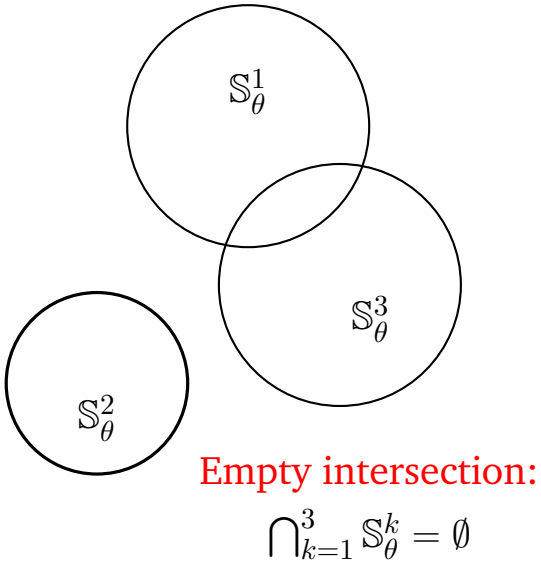


Figure 2.4 – Empty intersection of solution sets

Here, Equation (2.4) can be solved analytically, and we obtain

$$S_\theta^1 = [\underline{\mathbf{P}}, \overline{\mathbf{P}}] = [9090.90\ , 11111.11]$$

both the lower and upper bounds respectively obtained by considering $y = \overline{\tilde{y}}(= 0.0110)$ and $y = \underline{\tilde{y}}(= 0.009)$ in Equation (2.5). In the one dimensional case, the length of the interval, i.e., $\overline{\mathbf{P}} - \underline{\mathbf{P}}$ measures the area of $[\mathbf{P}]$, $\mathcal{A}([\mathbf{P}])$.

In chapter 3, we will consider a less illustrative example that is a typical mechanical inverse problem. It will consider the case where θ corresponds to the elastic Lamé parameters (λ and μ) and $\tilde{y}$ corresponds to full-field displacement data obtained after applying a given load, corresponding to the experimental conditions, on a material specimen.

In the general case, with multiple measurements, if all measurements are consistent, that is in our case if all intervals $[\underline{\tilde{y}_k}, \overline{\tilde{y}_k}]$ include the true value of the measured quantity y_k, then the solution set S_θ will be non-empty, as all S_θ^k in Equation (2.3) will have a non-empty intersection. However, it is very likely that some measurements will not contain this true value, and that they will be globally inconsistent. In such a case, we will have $\mathbb{S}_\theta = \emptyset$ as $\bigcap_{k=1}^{N} \mathbb{S}_\theta^k = \emptyset$ as illustrated in Figure 2.4. There may be several reasons for the inconsistency between the measurements and the model, such as the presence of measurement outliers or model error. Section 2.3 will present our solution to solve this issue, but before that we will recall the main solutions proposed in the literature, and the reasons why they may be unsatisfying in some cases.

2.2.2 Dealing with inconsistent measurements: issues with standard methods

In this section, we review existing methods to deal with the case of inconsistent measurements, that is when the exact inverse problem solution is the empty set, due to measurement uncertainty.

2.2.2.1 Average error minimization

A first way to solve the inverse problem in the presence of measurement uncertainty is to find the value θ^* minimizing the discrepancy between the reconstructed data $f(\theta) = y$ and the observed measurements $\tilde{y}$. Such techniques(least-squares) are discussed in the Section 1.2, and also, these methods have been extended to interval as well as to fuzzy approaches for identification problems, where one usually estimate first the central value of the interval (a.k.a. midpoint), before fixing its length (a.k.a. radius). We refer to [Faes and Moens, 2019, Section 5] for a short and relevant review of such techniques. In the case of intervals, they provide as answer an interval valued parameter $[\theta]^*$. These methods can be very sensitive to the presence of outliers. It should also be noted that nothing in their basic principles ensure that they will be consistent with existing measurements, in the sense that the result will have a non-empty intersection with some $\mathbb{S}_\theta^k$.

2.2.2.2 Bayesian inference method

We discussed the Bayesian inference method in Section 1.3. We refer to [Dashti and Stuart, 2017] for a full, recent study of such methods. Bayesian methods can be very efficient and accurate if the model is right. However, as for the least-squares methods, Bayesian methods can be quite sensitive to outliers (as we shall see in the experiments of Chapter 3), and require a strong modeling effort to be accurate and not be subject to model misspecification biases [Müller, 2013]. In this sense, their robustness can be limited, even if they are the first choice when it comes to model and process measurement uncertainty.

2.2.2.3 Q-intersection method

The fact that multiple set estimations can be inconsistent and have an empty intersection, as can happen in Equation (2.3), is a well-known fact in interval analysis. A common method to solve such issues is to consider the so-called q-intersection method [Sandretto et al., 2014, Drevelle and Bonnifait, 2012]. Given a set $\mathcal{S} = \{\mathbb{S}^1, \dots, \mathbb{S}^n\}$ of set-valued solutions, their q-intersection consists in applying

the formula

$$\overset{q}{\mathbb{S}} = \bigcup_{E\subseteq\mathcal{S},|E|=q} \bigcap_{\mathbb{S}_i\in E} \mathbb{S}_i, \tag{2.6}$$

that is to take the union over all possible intersections of q elements. While such an approach brings good results if the number n is not too large, and if q is close to n, it becomes quickly intractable when the number n of sets to merge increases, as the complexity of applying Equation (2.6) increases factorially with the number $n-q$.

Therefore, in the setting of inverse mechanical problems where the minimal number of measurements is some dozens and where it is not unusual to have thousands of measurements (e.g. digital image correlation data), the q-intersection method is not applicable. However, the methods proposed in the next Section 2.3.3 could be seen as a proxy of such an approach, where one searches to identify through heuristic methods a maximal subset of measurements within $\mathcal{S}$ whose intersection is non-empty.

2.2.2.4 Data consistency through relaxation

The method developed in [Hegde et al., 2018] and [Feeley et al., 2004] and henceforth called DC (for data consistency), applied to reaction modelling in chemistry, is very close in purpose to ours, as its intent is to quantify the inconsistency of a set of measurements and to provide a measurement-wise quantification of inconsistency. The method works as follows: let us denote by

$$\mathbb{S}_\theta^{k\pm\delta} = \{\theta \in \mathbb{S}_\theta^0 \mid f(\theta) \in [\underline{\tilde{y}_k} - \delta,\ \overline{\tilde{y}_k} + \delta]\} \tag{2.7}$$

the solution of the inverse problem when the kth measurement is broadened by a value δ. [Hegde et al., 2018] and [Feeley et al., 2004] propose to evaluate a data set inconsistency as

$$\delta_0 = \min\ \{\delta \in \mathbb{R} \mid \bigcap_{k=1}^{N} \mathbb{S}_\theta^{k\pm\delta} \neq \emptyset\}$$

which is negative if the measurements are consistent, and all the more positive as they become inconsistent. A way to define measurement-wise inconsistency is then provided through the use of Lagrange multipliers. The agenda of this method is very close to ours, yet we can notice some key differences:

- the DC method assumes that interval bounds are to a large extent subjective and modifiable, and search for a minimal correction making them consistent, while our approach does not modify the provided interval bounds;
- DC provides a refined consistency analysis and evaluation of measurement

inconsistency, at the expense of a costly optimization that lead to remove measurements with surgical precision. In contrast, our approach use coarser but easier to compute and handle evaluations, that are instrumental when one wants to solve the inverse problem by primarily (but not necessarily only) removing the most inconsistent measurements;

- while DC can be used to solve the inverse problem (i.e., provide a set of possible parameters), its primary goal is to finely detect inconsistencies between the measurements and the model. Conversely, while our approach can be used to analyse inconsistencies between the measurements and the model, its primary goal is to offer a solution to the inverse problem and propose a final set of parameter values.

While both methods rely on intervals to model uncertainty and consider similar notions of inconsistencies, they measure it differently and, most importantly, pursue different primary agendas.

2.3 Approaches to select sets of consistent measurements

We now describe our approach to tackle the presence of inconsistent measurements in the set-valued treatment of the inverse problem. Basically, our approach will aim at finding an important subgroup of consistent measurements by removing those measurements that are the most inconsistent with most of others.

This is different from the Bayesian and least-square approaches, as these methods try to find a compromise between all measurements, and do not discard any of them. Our methods are indeed closer to the q-intersection method, but do not consider every possible subgroup of q measurements, for reasons we just mentioned.

We first define how to assess the consistency of measurements (Sections 2.3.1 and 2.3.2) and then propose different outlier detection methods (Section 2.3.3) in the context of the identification strategy outlined in Section 2.2.1. We also discuss in Section 2.4 the implementation of the identification strategy in terms of a discrete description of sets to solve the set-valued inverse problems.

2.3.1 Pairwise measures of consistency

We first discuss measures of consistency between pairs of measurements, as they are intuitively easier to explain, and are instrumental to build more global consistency

measures.

2.3.1.1 Pairwise degree of inclusion

For any two solution sets $\mathbb{S}_\theta^k$ and $\mathbb{S}_\theta^{k'}$ corresponding respectively to $[\tilde{y}_k]$ and $[\tilde{y}_{k'}]$ measurements, $(k, k') \in \{1, ..., N\}^2$, we define the degree of inclusion (DOI) of one solution set $\mathbb{S}_\theta^k$ with respect to another $\mathbb{S}_\theta^{k'}$ as

$$DOI_{kk'} = \frac{\mathcal{A}(\mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'})}{\mathcal{A}(\mathbb{S}_\theta^{k'})} \tag{2.8}$$

where $\mathcal{A}(\mathbb{S})$ denotes the area of the set. The DOI between two solution sets is non-symmetric, i.e., $DOI_{kk'} \neq DOI_{k'k}$, and reaches its boundary values in the following situations

$$DOI_{kk'} = \begin{cases} 1 & \text{iff } \mathbb{S}_\theta^{k'} \subseteq \mathbb{S}_\theta^k \\ 0 & \text{iff } \mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'} = \emptyset \end{cases} \tag{2.9}$$

that are illustrated in Figure 2.5. The larger the value of DOI between one solution set and another, the higher the possibility of $\mathbb{S}_\theta^k$ included in $\mathbb{S}_\theta^{k'}$.

Example 2. Pursuing Example 1, we consider this time three measurements when a force F=100 N is applied on the spring, resulting in $\tilde{y}_1$=0.01 m, $\tilde{y}_2$=0.0101 m, and $\tilde{y}_3$=0.0110 m.

Assuming that our uncertainty on measurements is described by $[\underline{\tilde{y}_1}, \overline{\tilde{y}_1}] = [0.009\ , 0.0110]$, $[\underline{\tilde{y}_2}, \overline{\tilde{y}_2}] = [0.0091\ , 0.0111]$, $[\underline{\tilde{y}_3}, \overline{\tilde{y}_3}] = [0.01\ , 0.0120]$, we obtain by inversion of Equation (2.5) the estimations $[\mathbf{P}_1] = [9090.90\ , 11111.11]$, $[\mathbf{P}_2] = [9009\ , 10989]$, $[\mathbf{P}_3] = [8333.33\ , 10000]$.

As they have a non-empty intersection, they result in the final estimation $[\mathbf{P}] = [9090.90\ , 10000]$. They are such that $\mathcal{A}([\mathbf{P}_1]) = 2020.20$, $\mathcal{A}([\mathbf{P}_2]) = 1980.00$, $\mathcal{A}([\mathbf{P}_3]) = 1666.70$, $\mathcal{A}([\mathbf{P}_1] \cap [\mathbf{P}_2]) = 1898.10$, $\mathcal{A}([\mathbf{P}_1] \cap [\mathbf{P}_3]) = 909.10$, $\mathcal{A}([\mathbf{P}_2] \cap [\mathbf{P}_3]) = 991$, hence by knowing this information one can compute degree of inclusion between two measurements i.e., $DOI_{12} = 0.9586$, $DOI_{21} = 0.9396$, $DOI_{13} = 0.5454$, $DOI_{31} = 0.4500$, $DOI_{23} = 0.5946$, $DOI_{32} = 0.5005$, respectively.

2.3.1.2 Pairwise degree of consistency

We also define the degree of consistency (DOC) between two solution sets corresponding to $\mathbb{S}_\theta^k$ and $\mathbb{S}_\theta^{k'}$, which measures to which extent they agree with one another, considering that they fully agree (or are logically consistent) when one of them is totally included in the other. Indeed, such a situation means that they are in accordance, but that one is just more precise than the other. The following index

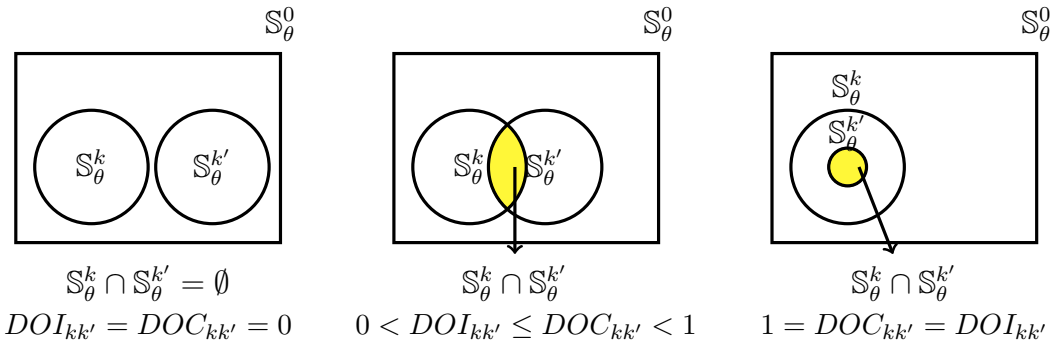


Figure 2.5 – DOI and DOC between two sets

translates this idea:

$$DOC_{kk'} = \frac{\mathcal{A}(\mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'})}{min(\mathcal{A}(\mathbb{S}_\theta^k), \mathcal{A}(\mathbb{S}_\theta^{k'}))} = \max(DOI_{kk'}, DOI_{k'k}). \tag{2.10}$$

It is symmetric ($DOC_{kk'} = DOC_{k'k}$) unlike the degree of inclusion, and reaches its boundary values in the following situations

$$DOC_{kk'} = \begin{cases} 1 & \text{iff } \mathbb{S}_\theta^k \subseteq \mathbb{S}_\theta^{k'} \text{ or } \mathbb{S}_\theta^{k'} \subseteq \mathbb{S}_\theta^k \\ 0 & \text{iff } \mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'} = \emptyset \end{cases} \tag{2.11}$$

that are illustrated in Figure 2.5.

Example 3. Using the numbers of Example 2, we have $DOC_{12} = 0.9586$, $DOC_{13} = 0.5454$, $DOC_{23} = 0.5946$. These values reflect how much each pair of measurements agree with one another, pairwisely.

2.3.2 Global measures of consistency

Pairwise measurements are interesting as they are easy to compute and limited in numbers, being equal to $N(N-1)$ if N is the number of measurements. However, focusing only on pairwise measurements may limit our perception of the problem, as we may well have non-null pairwise agreement between all sets, while having a global inconsistency between those sets.

For instance, the three sets $\mathbb{S}_\theta^1 = [3, 7]$, $\mathbb{S}_\theta^2 = [6, 10]$ and $\mathbb{S}_\theta^3 = [2, 4] \cup [8, 11]$ all have pairwise non-empty intersections, but a global inconsistency as $\cap_i \mathbb{S}_\theta^i = \emptyset$. In the N-dimensional case, such a situation is even possible with convex sets. For this reason, we also consider global measures of consistency for a set of measurements, that we introduce now. By global measures, we understand measures that consider all measurements at once, be it directly in their definition or by aggregating different pairwise measurements. Note that a particular case occurs if for a given k we have

$\mathbb{S}_\theta^k = \emptyset$, yet we will not consider this case, as we can just discard measurements that are inconsistent with the model.

2.3.2.1 Global degree of consistency (GDOC) of a measurement

We now introduce a measurement-wise consistency degree from a set of measurements. By using the pairwise degree of inclusion (DOI) of the solution sets corresponding to the measurements, we define the global degree of consistency ($GDOC$) of any k^{th} measurement with respect to all other measurements as

$$GDOC(k) = \frac{\sum_{k'=1}^{N} \frac{\mathcal{A}(\mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'})}{\mathcal{A}(\mathbb{S}_\theta^k)} + \sum_{k'=1}^{N} \frac{\mathcal{A}(\mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'})}{\mathcal{A}(\mathbb{S}_\theta^{k'})}}{2N} = \frac{\sum_{k'=1}^{N} (DOI_{kk'} + DOI_{k'k})}{2N}, \tag{2.12}$$

which reaches its boundary values in the following situations

$$GDOC(k) = \begin{cases} 1 & \text{iff } \mathbb{S}_\theta^1 = \mathbb{S}_\theta^2, ..., = \mathbb{S}_\theta^N \\ 0 & \text{iff } \mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'} = \emptyset, \quad \forall\, k' \in \{1, ..., N\} \end{cases} \tag{2.13}$$

The value of $GDOC(k)$ will always be between 0 and 1. It should be noted that the conditions imposed on this index to reach its boundary values are very strong, as having $GDOC = 1$ for one measurement requires all sets to be absolutely identical, the strongest possible consistency between sets. On the other side, $GDOC(k) = 0$ corresponds to the case where the k^{th} measurement is a complete outlier, i.e., has empty intersection with all other sets. Therefore, a high value of $GDOC$ for the k^{th} measurement indicates that it has a high consistency with the other measurements. $GDOC$ is therefore a good measure of the consistency of a single measurement, and can be used as a selection method, as suggested in Section 2.3.3.1 and Algorithm 2. Nevertheless, even if the measure accounts for all measurements, it makes sense to complement its measurement-centered view by a group-wise measure.

Example 4. Using results from Example 1, we can compute GDOC of each measurement i.e., $GDOC(1) = 0.8156$, $GDOC(2) = 0.8322$, $GDOC(3) = 0.6818$. These values provide the consistency of individual measurements.

2.3.2.2 General measure of consistency

Note that while GDOC is a global measure, it may still suffer from the drawbacks illustrated in the example given at the beginning of Section 2.3.2. So, we now define a global consistency measure for a group of measurements. Let $S = \{\mathbb{S}_\theta^1, \ldots, \mathbb{S}_\theta^k, \ldots, \mathbb{S}_\theta^N\}$ be the set of solutions to the inverse problems for the measurements $\{\tilde{y}_1,..\tilde{y}_k...,\tilde{y}_N\}$, with $\mathbb{S}_\theta^k \subseteq \mathbb{R}^M$. We define the general consistency

($GCONS$) for any subset $E \subseteq S$ of solution sets corresponding to a subset of the measurements as

$$GCONS(E) = \frac{\mathcal{A}(\bigcap_{\mathbb{S}_\theta^k \in E} \mathbb{S}_\theta^k)}{\min\limits_{\mathbb{S}_\theta^k \in E} \mathcal{A}(\mathbb{S}_\theta^k)} \tag{2.14}$$

It has the following properties:

1. It is insensitive to permutation of the sets of measurement (commutativity).
2. The value of $GCONS$ is monotonically decreasing with the size of the set E, in the sense that for any subsets E and F, with $E \subseteq F$, we have $GCONS(F) \leq GCONS(E)$. It also means that the more measurements we have, the less consistent they can be with one another.
3. The boundary values of equation (2.14) are obtained for the following situations

$$GCONS(E) = \begin{cases} 1 & \text{iff } \mathcal{A}(\bigcap\limits_{\mathbb{S}_\theta^k \in E} \mathbb{S}_\theta^k) = \min\limits_{\mathbb{S}_\theta^k \in E} \mathcal{A}(\mathbb{S}_\theta^k) \\ 0 & \text{iff } \mathcal{A}(\bigcap\limits_{\mathbb{S}_\theta^k \in E} \mathbb{S}_\theta^k) = \emptyset \end{cases}$$

   where the first case means that at least one measurement is fully consistent (i.e., included) with all the other measurements corresponding to E, that therefore have a non-empty intersection between them.
4. When $|E| = 2$, $GCONS(E)$ boils down to Equation (2.10), therefore effectively generalizing DOC to any subset of solution sets, and not only pairs.

Example 5. Using the numbers of Example 2, we can compute $GCONS$ of the set of three measurements i.e., $GCONS(E) = 0.5454$ using Equation (2.14). From this, we can already see that all measurements taken together are reasonably consistent with each others.

2.3.3 Algorithmic procedures

We propose three algorithms describing different outlier detection methods, which make use of the $GDOC$, the pairwise degree of consistency DOC and the $GCONS$ measure, respectively. The ideas of these algorithms is that they are efficient heuristics to retrieve a set $E \subseteq S$ corresponding to consistent measurements, that is they aim at removing a small number of measurements while ensuring a high consistency of the remaining measurements, the goal being to have an informative and reliable solution to the inverse problem.

2.3.3.1 Individual approach

The first approach we will describe relies on the individual measures of global consistency based on $GDOC$. To describe it, let us first consider the permutation $\pi : \{1, \ldots, N\} \to \{1, \ldots, N\}$ such that $GDOC(\pi(i)) \geq GDOC(\pi(j))$ if $i < j$, where $\pi(k)$ is the index of the measurement within $\{[\tilde{y}_1], \ldots, [\tilde{y}_N]\}$ that has the k^{th} biggest $GDOC$ measure. That is, we rank the measurements from the most globally consistent to the least consistent. We now introduce the function $f_{GDOC} : [1, N] \to [0, 1]$ such that for each $k \in \{1, \ldots, N\}$ we have

$$f_{GDOC}(k) = GDOC(\pi(k)), \tag{2.15}$$

and a linear interpolation between any two of those points (used to detect and better visualize the inflection point as detailed below). Such a decreasing function is plotted in Figure 3.5. The idea of our first method, described in Algorithm 2, is straightforward: removing measurements with the lowest $GDOC$ values, until the increase of the minimal $GDOC$ of the remaining measurements after one suppression is negligible and it is based on the bending of the curve (see Figure 3.5). Note that in those cases where an inflection point cannot be found, one can always set a pre-defined threshold τ in line 8 of Algorithm 2. However, we never witnessed this situation in our experiments.

Algorithm 2: GDOC method

Input: Set $S = \{\mathbb{S}_\theta^1, \ldots, \mathbb{S}_\theta^N\}$ of inverse problem solutions
Output: Subset E of selected measurements

1 **begin**
2 $E \leftarrow \emptyset$;
3 **for** *all pairs* $[\tilde{y}_i], [\tilde{y}_j]$ *of measurements* **do**
4 Compute DOI_{ij}, DOI_{ji} ;
5 **for** $1 \leq k \leq N$ **do**
6 Compute $GDOC(k)$ from Equation (2.12) ;
7 Detect point x^* of inflection of f_{GDOC} ;
8 $\tau \leftarrow f_{GDOC}(x^*)$;
9 **return** $E \leftarrow \{\mathbb{S}_\theta^i \in S : GDOC(i) > \tau\}$;

2.3.3.2 Pairwise approach

While the previous approach focused on removing the measurements that were globally the most inconsistent with the others, it did not look at potential interaction of such measurements, in the sense that once some measurements are removed, a

measurement at first deemed poorly consistent could well be consistent with the remaining ones.

An easy way to capture such potential interaction is through pairs of measurements, and our second method follows this idea by using the degree of consistency (DOC) between two solution sets corresponding to k^{th} and k'^{th} measurement given by Equation (2.10). To handle them, we first build a $N \times N$ matrix C as follows:

$$C = \begin{bmatrix} DOC_{11} & DOC_{12} & .. & DOC_{1N} \\ DOC_{21} & DOC_{22} & .. & DOC_{2N} \\ .. & .. & .. & .. \\ DOC_{N1} & DOC_{N2} & .. & DOC_{NN} \end{bmatrix} \tag{2.16}$$

It is easy to see that this is a symmetric matrix ($C_{kk'} = C_{k'k}$), as DOC is a symmetric measure, that its diagonal elements always[1] have value 1, and the off-diagonal elements are all within $[0, 1]$.

Given the characteristic of matrix C, any k^{th} row or column of the matrix represents the pairwise DOC values of the k^{th} measurement. Basically, each k^{th} row provides some idea of the overall consistency of the k^{th} measurement with respect to other measurements. This is similar to $GDOC$, yet we abstain to aggregate such values, thus preserving more details about the measurement interactions.

To detect which measurements should be removed according to this pairwise information, we will associate to each measurement a consistency count, that we will denote by $Count : \{1, \ldots, N\} \rightarrow \mathbb{N}$, with $Count(k)$ the count associated to the k^{th} measurement. Basically, the idea behind those counts is to assess whether removing a measurement from matrix C could significantly increase the minimal value of pairwise consistencies. This is also why, in practice, we will focus on small pairwise consistencies, as we are not interested in removing measurements quite consistent with all the others. Algorithm 3 summarises the procedure:

- Line 4 in the algorithm limits the search to the measurements having the highest pairwise inconsistency.
- Loop 7 to 12 starts from the most inconsistent pair, and then looks which of the two involved measurements (i_k and j_k) is also the most inconsistent with respect to other measurements (using standard notations, we denote by $C_{i_k:}$ and $C_{:j_k}$ the i_kth line and j_kth column of C).
- Lines 13 and 14 simply ensure that if index i_k or j_k appears in a further

[1] Unless one has an inconsistent measurement $\mathbb{S}^k_\theta = \emptyset$, but such measurement should be removed before starting any further analysis, so we can assume that we have none.

iterations, then the minimum in lines 7 to 10 will not be obtained for $C_{i_k j_k}$, avoiding some double counting.

Algorithm 3: Count method

Input: C matrix, threshold τ of minimal DOC value
Output: Function $Count$

1 **begin**
2 **for** $k \in \{1, \ldots, N\}$ **do**
3 $Count(k) \leftarrow 0$
4 $\mathcal{C}_\tau \leftarrow \{C_{ij} : C_{ij} < \tau\}$;
5 Sort $\mathcal{C}_\tau$ such that $C_{i_1 j_1} \leq \ldots \leq C_{i_{|\mathcal{C}_\tau|} j_{|\mathcal{C}_\tau|}}$;
6 **for** $k \in \{1, \ldots, |\mathcal{C}_\tau|\}$ **do**
7 **if** $(\min\{C_{i_k:} \setminus C_{i_k j_k}\} > \min\{C_{:j_k} \setminus C_{i_k j_k}\}$ **then** $Count(j_k) \leftarrow Count(j_k) + 1$;
8 **if** $(\min\{C_{i_k:} \setminus C_{i_k j_k}\} < \min\{C_{:j_k} \setminus C_{i_k j_k}\}$ **then**
9 $Count(i_k) \leftarrow Count(i_k) + 1$
10 **if** $(\min\{C_{i_k:} \setminus C_{i_k j_k}\} = \min\{C_{:j_k} \setminus C_{i_k j_k}\}$ **then**
11 $Count(i_k) \leftarrow Count(i_k) + 1$;
12 $Count(j_k) \leftarrow Count(j_k) + 1$;
13 $C_{i_k j_k} \leftarrow 1$;
14 $C_{j_k i_k} \leftarrow 1$;
15 **return** $Count$;

Example 6. We consider the following illustrative C matrix, collected from 4 measurements.

$$C = \begin{bmatrix} 1 & 0.7 & 0.8 & 0.05 \\ 0.7 & 1 & 0.6 & 0.2 \\ 0.8 & 0.6 & 1 & 0.1 \\ 0.05 & 0.2 & 0.1 & 1 \end{bmatrix}$$

From Matrix C, we can observe that the fourth measurement has a low DOC value with the first, second, and third measurements, and is therefore a likely outlier or bad measurement. Applying Algorithm 3 to this matrix with a threshold of $\tau = 0.15$, we get a set[2] $\mathcal{C} = \{C_{14}, C_{34}\}$, with $i_1 = 1$, $j_1 = 4$ and $i_2 = 3$, $j_2 = 4$. The first application of Loops 7-14 in the Algorithm shows that $\min\{C_{1:} \setminus C_{14}\} = 0.7 > \min\{C_{:4} \setminus C_{14}\} = 0.1$ and we therefore increment $Count(4)$ from 0 to 1. Finally, after the second loop, we get $Count(4) = 2$ and $Count(k) = 0$ for all other $k \in \{1, 2, 3\}$.

Our next step is to determine which measurements should be removed according to our count function. Similarly to what was done in Section 2.3.3.1, let us consider

[2]Note that we can focus only on the upper/lower triangle part of the matrix, as it is symmetric.

the permutation $\pi : \{1, \ldots, N\} \to \{1, \ldots, N\}$ such that $Count(\pi(i)) \geq Count(\pi(j))$ if $i < j$, where $\pi(k)$ is the index of the measurements within $\{[\tilde{y}_1], \ldots, [\tilde{y}_N]\}$ that has the k^{th} biggest count. That is, we rank the measurements from the one with the most count (i.e., the most inconsistent according to the second method) to the one with the fewest. We now introduce the function $f_{Count} : [1, N] \to [0, 1]$ such that for each $k \in \{1, \ldots, N\}$ we have

$$f_{Count}(k) = Count(\pi(k)), \tag{2.17}$$

and a linear interpolation between any two of those points. Such a decreasing function is plotted in Figure 3.7(b). We then apply the same strategy as before, except that this time we retain all measurements whose count is lower than the one corresponding to f_{Count} inflection point. This is summarised in Algorithm 4.

Algorithm 4: Count selection method

Input: Set $S = \{\mathbb{S}_\theta^1, \ldots, \mathbb{S}_\theta^N\}$ of inverse problem solutions, f_{Count}
Output: Subset E of selected measurements

begin
  Detect point x^* of inflection of f_{Count} ;
  $\tau \leftarrow f_{Count}(x^*)$;
  return $E \leftarrow \{\mathbb{S}_\theta^i \in S : Count(i) < \tau\}$;

2.3.3.3 Global approach

Algorithm 5: $GCONS$ selection method

Input: Sorted set $S = \{\mathbb{S}_\theta^1, \ldots, \mathbb{S}_\theta^N\}$ such that
$GDOC(1) \geq \ldots \geq GDOC(N)$, threshold τ
Output: Subset E of solution sets corresponding to the selected measurements

begin
  $E \leftarrow \{\mathbb{S}_\theta^1, \mathbb{S}_\theta^2\}$; `// Initialization`
  for $k \in \{3, \ldots, N\}$ **do**
    if $GCONS(E \cup \mathbb{S}_\theta^k) \geq \tau$ **then** $E \leftarrow E \cup \mathbb{S}_\theta^k$;
  return E;

Our third method considers global consistency information rather than more localized ones, and rely on the properties of $GCONS$ introduced in the Section 2.3.2.2. In theory, a good principle to choose a subset of consistent measurements would be to search for the largest subset $E \subseteq S$ (corresponding to the maximal number of measurements) that has a reasonable consistency, that

is for which $GCONS(E)$ is above some threshold τ. Yet, such a search could be at worst exponential in N, which as recalled in Section 2.2.2 can be quite large in mechanical problems, where it is not rare to have thousands, if not tens of thousands of measurements. Such an approach would therefore be untractable. This is why we propose a greedy algorithm (Algorithm 5) that makes use of $GDOC$ measures to find a suitable subset E. The idea is again quite straightforward: starting from the most consistent solution set corresponding to the most consistent measurement according to $GDOC$ and ordering them according to their individual consistency, we iteratively add new solution set corresponding to measurements to E unless they bring the global consistency $GCONS$ under a pre-defined threshold, that is unless they introduce too much inconsistency.

2.4 Implementation with discrete description of sets

Before applying any of the methods we just described, a pre-requisite is to obtain the sets $\mathbb{S}_\theta^1, \ldots, \mathbb{S}_\theta^N$, or a numerical approximation of them. There are multiple ways to obtain such approximations, such as using boxes and bi-section methods, of which SIVIA algorithm [Jaulin and Walter, 1993] is a well-known example, or using a grid of points for which the solution is known. The first solution has the advantage to provide guarantees, in the sense that the approximation is known to contain the true set, but can be computationally intensive, especially for high-dimensional problems. The second solution seems more fitted to our case, as having strong guarantees is not essential (as long as the approximation is sufficiently accurate), and as the number of measurements and the dimensions can be quite high, especially compared to robotics applications where the input dimensions is usually limited to 2 or 3 dimensions for localisation problems, and the number of measurements (issued from sensors or satellites) is usually at most around 10. In those cases where even using a uniform grid of points is prohibitive (due to many input or parameter dimensions or a costly model to run), one can resort to several different strategies to speed up the computations, such as the use of adaptive, non-uniform grids [Sui et al., 2018], the use of surrogate models [Fang et al., 2018, Xiaoguang et al., 2019] or of techniques dedicated to the application at hand when those exist.

Here, we use the same description as in [Sui et al., 2018], that is a grid of points, θ_i, $i \in \{1, ..., N_g\}$ as shown in Figure 2.6(a) where N_g is the number of grid points. Such a description, illustrated by Figure 2.6(b), is convenient when comparing or intersecting the sets since the grid of points is the same for all measurements and all sets. Any set $\mathbb{S}_\theta \subset \mathbb{S}_\theta^0$ is then characterized through its discrete characteristic

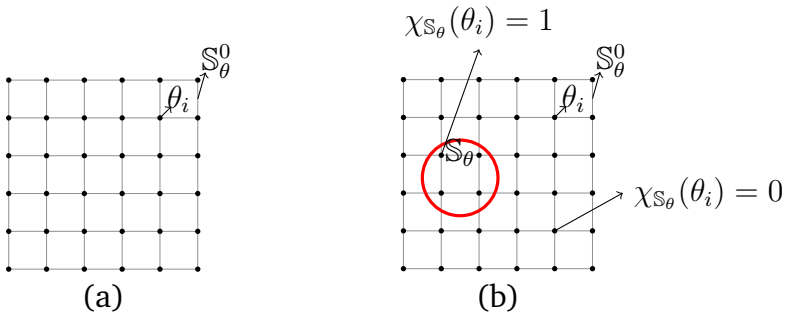


Figure 2.6 – Discrete description of sets. (a) Prior set ($\mathbb{S}_\theta^0$). (b) Characterized set ($\mathbb{S}_\theta$)

function, defined at any point $\theta_i \in \mathbb{S}_\theta^0$ of the grid as

$$\chi_{\mathbb{S}_\theta}(\theta_i) = \begin{cases} 1 & \text{if } \theta_i \in \mathbb{S}_\theta \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

In this work, a uniform grid is chosen to describe the prior parameter set $\mathbb{S}_\theta^0$, but it is not mandatory. In our method, each $\mathbb{S}_\theta^k$ is therefore described by its discrete characteristic function, defined at any point of the grid as

$$\chi_{\mathbb{S}_\theta^k}(\theta_i) = \begin{cases} 1 & \text{if } \underline{\tilde{y}_k} \leq f(\theta_i) \leq \overline{\tilde{y}_k} \\ 0 & \text{otherwise} \end{cases} \tag{2.19}$$

These discrete characteristic functions can be collected in a $N_g \times N$ matrix X as columns of boolean values as follows

$$X = \begin{bmatrix} 1 & 1 & .. & 1 \\ 0 & 1 & .. & 1 \\ .. & .. & .. & .. \\ 1 & 1 & .. & 0 \end{bmatrix} \tag{2.20}$$

where $\chi_{\mathbb{S}_\theta^k}(\theta_i)$ is the element of column k and line i. Using matrix X, a $N \times N$ symmetric matrix $T = X^T X$ can be obtained, whose components are directly proportional to the areas of the sets $\mathbb{S}_\theta^i$, and can therefore be used as an estimation of such areas, i.e.,

$$T \propto \begin{bmatrix} \mathcal{A}(\mathbb{S}_\theta^1) & \mathcal{A}(\mathbb{S}_\theta^1 \cap \mathbb{S}_\theta^2) & .. & \mathcal{A}(\mathbb{S}_\theta^1 \cap \mathbb{S}_\theta^N) \\ \mathcal{A}(\mathbb{S}_\theta^2 \cap \mathbb{S}_\theta^1) & \mathcal{A}(\mathbb{S}_\theta^2) & .. & \mathcal{A}(\mathbb{S}_\theta^2 \cap \mathbb{S}_\theta^N) \\ .. & .. & .. & .. \\ \mathcal{A}(\mathbb{S}_\theta^N \cap \mathbb{S}_\theta^1) & \mathcal{A}(\mathbb{S}_\theta^N \cap \mathbb{S}_\theta^2) & .. & \mathcal{A}(\mathbb{S}_\theta^N) \end{bmatrix} \tag{2.21}$$

Indeed, the diagonal element T_{kk} of T represents the number of grid points that are in the solution set of the inverse problem for the k^{th} measurement, and it is proportional to $\mathcal{A}(\mathbb{S}_\theta^k)$. The non-diagonal element $T_{kk'}$ of T represents the number of grid points for which both kth and k'th measurements are consistent and it is proportional to $\mathcal{A}(\mathbb{S}_\theta^k \cap \mathbb{S}_\theta^{k'})$. Hence, $GDOC$ can be computed from matrix T for any k^{th} measurement as

$$GDOC(k) = \frac{\sum_{k'=1}^{N} \frac{T_{k'k}}{T_{kk}} + \sum_{k'=1}^{N} \frac{T_{kk'}}{T_{k'k'}}}{2N}. \tag{2.22}$$

2.5 Summary

In this Chapter, we have presented a new parameter identification strategy relying on set theory and on interval measurements. In this approach, we have used intervals to describe uncertainty on measurements and parameters. We have presented different consistency measures for measurements and used them to propose three measurement selection methods taking advantage of $GDOC$, $Count$ and $GCONS$. We have also discussed the implementation of the identification strategy using discrete description of sets and how such a description of the sets helps to find the area of the solution set. The next Chapter will be devoted to the application of the identification strategies to an identification problem of elastic properties. We will in particular focus on the problem of identifying outliers, and will compare our approaches with the standard Bayesian approach, showing that in the presence of outliers, our approaches indeed provide more reliable results, while we remain competitive with the Bayesian method in the ideal situation where the assumptions of the Bayesian model are verified.

Dealing with inconsistent measurements in inverse problems

Contents

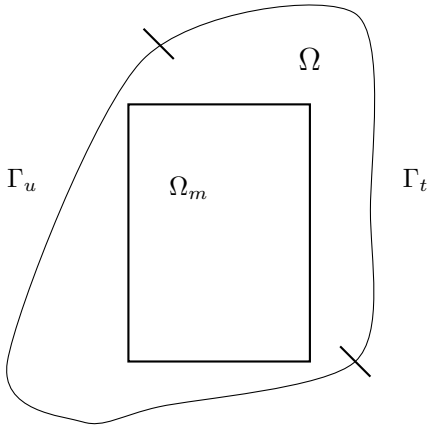


Figure 3.1 – Mechanical model for identification

3.1 Introduction

This chapter focuses on applying a set-valued inverse method to obtain a solution to the mechanical inverse problem, mainly when measurements are inconsistent with one another. The mechanical inverse problem we have considered consists of identifying mechanical properties of materials based on parametrized mechanical models. We try to identify the mechanical properties of a homogeneous isotropic elastic structure using the full-field displacement measurements. In this work, the full-field displacement measurements are assumed to be obtained using digital image correlation (DIC) from CCD camera images [Grédiac and Hild, 2012]. We also illustrate three selection (outlier detection) methods to deal with inconsistent measurements and compare the set-valued inverse method to the Bayesian inference method in situations where the model is well specified and where it is not.

3.2 Mechanical model

This section discusses the mechanical model (elastic structure) using continuum mechanics equations that also describe the experiment's modelling. Then, we briefly introduce the Finite Element Method (FEM) used to discretize continuum mechanics equations.

3.2.1 Mechanical model

Continuum solid mechanics defines the behaviour of solid materials, e.g., motion and deformation under external loadings. Mechanical models usually contain

parameters characterizing mechanical systems (e.g., loadings, geometry, mechanical property such as elasticity, elastoplasticity, etc.) and their relations. In this work, we aim at the identification of elastic material properties. Thus, we only describe the linear isotropic elastic mechanical model. We consider the solid structure shown in Figure 3.1, which describes a 2D domain Ω. The model equations that govern the behaviour of the structure follow the below assumptions

- deformations are small;
- the behaviour of the material is linear;
- dynamic effects are neglected;
- the material is homogeneous.

The governing equations are as follows:

- on Ω:
  Equilibrium equation:
  $$\text{div}\boldsymbol{\sigma} = 0, \tag{3.1}$$
  Kinematics equation (strain–displacement relation):
  $$\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla \boldsymbol{u} + \nabla^T \boldsymbol{u}), \tag{3.2}$$
  Constitutive equation (stress–strain relation):
  $$\boldsymbol{\sigma} = \boldsymbol{D}(\boldsymbol{\theta})\varepsilon \tag{3.3}$$
  where $\boldsymbol{u}$ is the displacementent field, $\boldsymbol{\sigma}$ is the stress field, ε is the strain field, $\boldsymbol{D}(\boldsymbol{\theta})$ is the elastic tensor field which can be described using parameters $\boldsymbol{\theta}$ that denoting the elastic properties. In this work, parameters $\boldsymbol{\theta}$ are λ and μ. They are termed as Lamé parameters and denote the elastic properties;
- on Γ_u: $\boldsymbol{u} = \boldsymbol{u_d}$, where $\boldsymbol{u_d}$ is a known Dirichlet boundary condition(here it is a displacement);
- on Γ_t: $\boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{t_n}$, where $\boldsymbol{t_n}$ is a known Neumann boundary condition(here it is a traction), and we also assume that $\Gamma_u \cup \Gamma_t = \Gamma$ and $\Gamma_u \cap \Gamma_t = \emptyset$;
- on Ω_m: $\boldsymbol{u} = \tilde{\boldsymbol{u}}$, where $\tilde{\boldsymbol{u}}$ is the displacement related to the DIC measurement on a sub-domain $\Omega_m \subset \Omega$.

From Equations (3.1) to (3.3) and the boundary conditions on Γ_u and Γ_t, a well-posed direct problem can be defined. Its solution is given as

$$\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{\theta}) \tag{3.4}$$

We perform the identification of parameters $\boldsymbol{\theta}$ by comparing $\boldsymbol{u}(\boldsymbol{\theta})$ with $\tilde{\boldsymbol{u}}$ on Ω_m which gives some information on $\boldsymbol{\theta}$.

3.2.2 Finite Element Method

The Finite Element Method is used to discretize Equation (3.1). The weak form of (3.1) is written as:

$$W_* = \int_\Omega \boldsymbol{u}_* \cdot \text{div}(\boldsymbol{\sigma}) d\Omega = 0, \ \forall \ \boldsymbol{u}_* \in H_0^1(\Omega), \tag{3.5}$$

where $\boldsymbol{u}_*$ denotes a virtual displacement field, $H_0^1(\Omega)$ is a Hilbert space, $\boldsymbol{\sigma}$ is the stress field which can be written in vector form (2D case):

$$\boldsymbol{\sigma} = (\sigma_{xx}, \ \sigma_{yy}, \ \sigma_{xy})^T$$

The strain field in the 2D case can be written as:

$$\boldsymbol{\varepsilon} = (\varepsilon_{xx}, \ \varepsilon_{yy}, \ \varepsilon_{xy})^T$$

Correspondingly, the elastic tensor $\boldsymbol{D}(\boldsymbol{\theta})$ in the 2D case is represented by a 3×3 matrix. In 2-D problems, the matrix $\boldsymbol{D}$ depends on whether one assumes a plane stress or a plane strain condition. In our work we choose plain strain condition and in that case for 2D isotropic linear elasticity problem it is given as:

$$\begin{bmatrix} \lambda + 2\mu & \lambda & 0 \\ \lambda & \lambda + 2\mu & 0 \\ 0 & 0 & \mu \end{bmatrix}$$

By applying the partial integration formula, we can transform Equation (3.5) into the following form:

$$W_* = \int_\Omega \boldsymbol{\varepsilon}_* \cdot \boldsymbol{\sigma} d\Omega + \int_{\Gamma_t} \boldsymbol{u}_* \cdot \boldsymbol{\sigma} \cdot \boldsymbol{n} d\Gamma, \ \forall \ \boldsymbol{u}_* \in H_0^1(\Omega), \tag{3.6}$$

where $\boldsymbol{\varepsilon}_*$ is the virtual strain field associated with the virtual displacement field $\boldsymbol{u}_*$. Then Equations (3.2) and (3.3) are taken into account to have a weak form in

terms of displacements. Considering $\boldsymbol{u_*}$ is null on Γ_u and considering the Neumann boundary condition, the boundary terms are written as:

$$\int_{\Gamma} \boldsymbol{u_*} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{n} d\Gamma = \int_{\Gamma_t} \boldsymbol{u_*} \cdot \boldsymbol{t_n} d\Gamma \tag{3.7}$$

To discretize Equation (3.6), the division of the domain into elementary polyhedrons (called Finite Element meshing) should be done firstly. Then the weak form can be discretized and appears as a sum of elementary terms

$$W_* = \sum_{e=1}^{nelt} (W_\Omega^e - W_{BC}^e) = 0, \tag{3.8}$$

with

$$W_\Omega^e = \int_{\Omega^e} \boldsymbol{\varepsilon_*} \cdot \boldsymbol{D(\theta)} \cdot \boldsymbol{\varepsilon} d\Omega \tag{3.9}$$

$$W_{BC}^e = \int_{\Gamma_t^e} \boldsymbol{u_*} \cdot \boldsymbol{t_n} d\Gamma, \tag{3.10}$$

where Ω is the integral domain, BC corresponds to the boundary conditions, Ω^e is the domain of a single finite element ($\Omega^e \subset \Omega$), Γ_t^e is the boundary of Ω, if any, which locates on the considered finite element). The number $nelt$ is the number of finite elements used to create the mesh. The displacement fields($\boldsymbol{u}$ and $\boldsymbol{u_*}$) of each element are represented by the displacements in each node and corresponding approximation:

$$\boldsymbol{u(x)} = \sum_{i=1}^{n_e} N_i(x) u_i, \tag{3.11}$$

where u_i is the displacement of the node i, $N_i(x)$ is the shape function corresponding to the node i and n_e is the number of nosed on element. The shape functions depend only on the coordinates of the elements nodes. By substituting Equation (3.11) into Equation (3.8), we can transform the weak form into the discrete matrix form as:

$$W_* = \underset{e=1}{\overset{nelt}{\boldsymbol{A}}} (\boldsymbol{U_*^e})^T (\boldsymbol{K^e U^e - F^e}) = 0, \ \forall \boldsymbol{U_*^e}, \tag{3.12}$$

where the symbol $\boldsymbol{A}$ is the operator that assembles the elementary matrices into a global matrix. The vector $\boldsymbol{U_*^e}$ collects the nodal components of the virtual displacement. The matrix $\boldsymbol{K^e}$ is the elementary matrix and $\boldsymbol{F^e}$ is the nodal force vector. They can be written as

$$\boldsymbol{K^e} = \int_{\Omega^e} \boldsymbol{B^{eT} D^e B^e} d\Omega, \tag{3.13}$$

$$\boldsymbol{F}^{\boldsymbol{e}} = \int_{\Gamma_t^e} N^{eT} t_n d\Gamma, \tag{3.14}$$

where $\boldsymbol{B}^e$ is the gradient matrix of the shape functions, $\boldsymbol{D}^e$ is the matrix associated with the material properties, N^e is the element shape function matrix. As $\boldsymbol{u}_*$ is totally independent from actual displacements, we can write Equation (3.12) as follows:

$$\underset{e=1}{\overset{nelt}{\boldsymbol{A}}} (\boldsymbol{K}^e \boldsymbol{U}^e - \boldsymbol{F}^e) = 0 \tag{3.15}$$

Assembling the elementary local matrices and vectors, we can obtain a global linear system of equations:

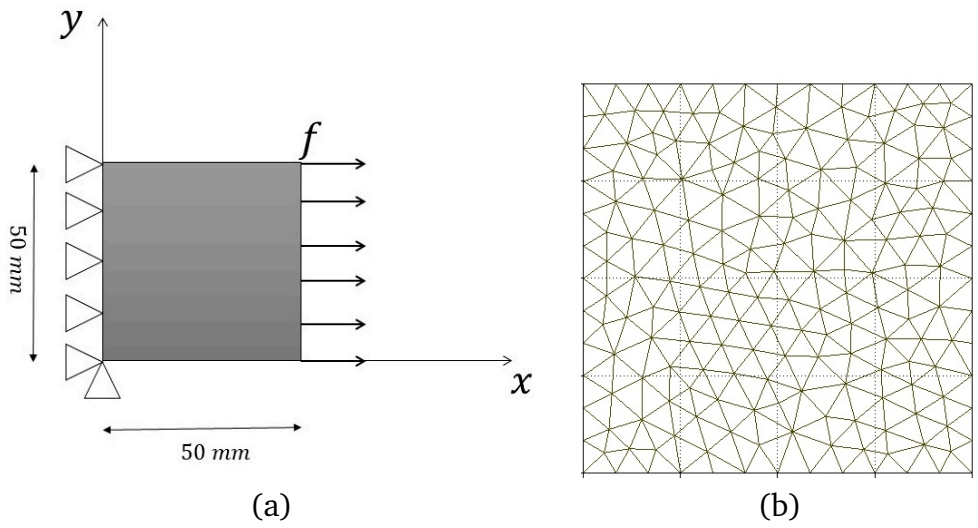
$$\boldsymbol{K}\boldsymbol{U} = \boldsymbol{F} \tag{3.16}$$

where $\boldsymbol{K}$ is the stiffness matrix with $2n \times 2n$ size, $\boldsymbol{U}$ is the vector of displacement with size $2n \times 1$ and $\boldsymbol{F}$ is the vector of external force with size $2n \times 1$. The value n is the total number of nodes. In the next section and Chapter, we will denote displacement with a notation y and full-field displacement measurement with a notation $\tilde{y}$. We use Equation (3.16) to compute the displacement y corresponding to the parameters (λ and μ) and external load $\boldsymbol{F}$.

3.3 Numerical applications

In this Section, we apply the set-valued inverse method to identify elastic properties (Lamé parameters: λ and μ) of a homogeneous 2D plate under plane strain as shown in Figure 3.2(a). The plate is clamped on the left side and loaded on the right side by a uniform traction $f = 1000\ N/m$. To generate displacement measurement data $\tilde{y}$ (386 measurements), exact displacement data y^{Ref} are simulated by a Finite Element (FE) model (193 nodes, 336 elements) as shown in Figure 3.2(b) considering the reference values $\lambda_0 = 1.15 \cdot 10^5 MPa$ and $\mu_0 = 7.69 \cdot 10^4 MPa$. We also consider a Gaussian noise with 0 mean (no systematic bias) and with standard deviation σ, taken here as 5% of the average of all the exact displacement values. In practical cases, it can be assumed that σ can be deduced from the measurement technique.

To apply our methods, the uncertainty on the measurements has to be given in interval form. Therefore, each measurement is modelled as $[\tilde{y}_k - 2\sigma, \tilde{y}_k + 2\sigma]$. The width of 2σ ensures that sufficient measurements will be consistent with one another. Prior information about the parameters ($\mathbb{S}_\theta^0$) is considered as a $2D$ box $\lambda^p \times \mu^p$ with $\lambda^p = [0.72 \cdot 10^5, 1.90 \cdot 10^5] MPa$ and $\mu^p = [\,7.2 \cdot 10^4, 8.15 \cdot 10^4] MPa$. All grids used in the experiments counts 3600 points (60×60). In the next section, we illustrate the effect of our approach in two situations: when data is noiseless,



(a) (b)

Figure 3.2 – A homogeneous plate and its model. (a) 2D homogeneous plate. (b) FEM mesh.

and when some noise is present in the data.

3.3.1 Identification with different outlier detection methods: illustration

Before proceeding to a statistical analysis of our different approaches, we first provide some illustrations of how they work, respectively for the $GDOC$, $Count$, and $GCONS$ selection methods presented in Chapter 2.

In each case, we consider two situations:

1. in the first one the measurement data are exact, i.e., $\tilde{y} = y^{Ref}$ and the information on the measurement $\tilde{y}$ was described in an interval form: $[\ \tilde{y} - 2\sigma,\quad \tilde{y} + 2\sigma\ ]$;

2. in the second one, we first add to each $\tilde{y}$ a random Gaussian white noise (with 0 mean) having a standard deviation of σ. The resulting value $\tilde{y}$ is then transformed into an interval $[\ \tilde{y} - 2\sigma,\quad \tilde{y} + 2\sigma\ ]$

3.3.1.1 $GDOC$ method

Figure 3.3 shows (in yellow) the intersection of the 386 computed sets $S_\theta^1, \ldots, S_\theta^N$ in absence of noise. Of course, since there are no errors, there is no inconsistency in the measurements and the feasible set contains the true parameter values. It can be noted that the absence of noise and the modeling of error by $[\ \tilde{y} - 2\sigma,\quad \tilde{y} + 2\sigma\ ]$ leads to an overestimation of the identification of a solution set.

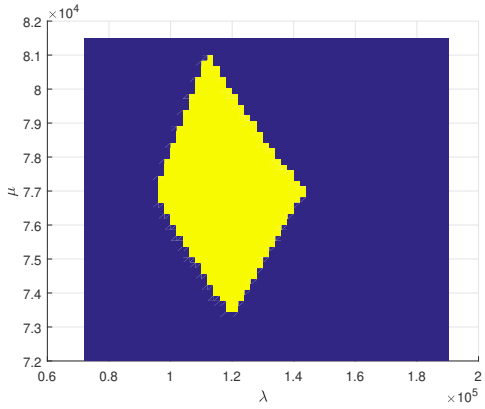
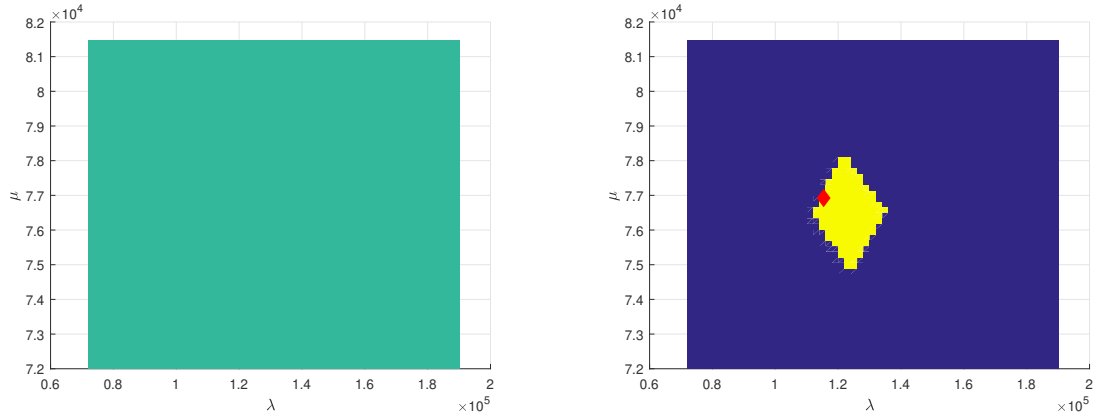


Figure 3.3 – Feasible set of parameters



(a) (b)

Figure 3.4 – Identification with $GDOC$ method. (a) Empty solution set (all measurements). (b) Solution set after selection.

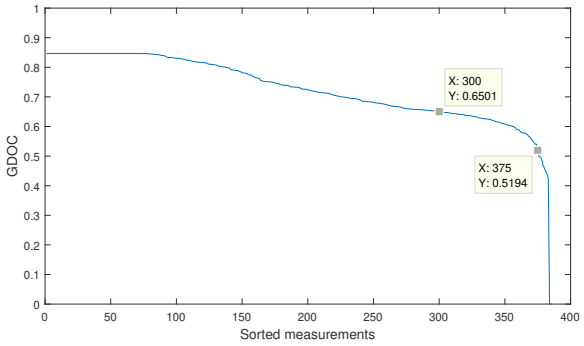
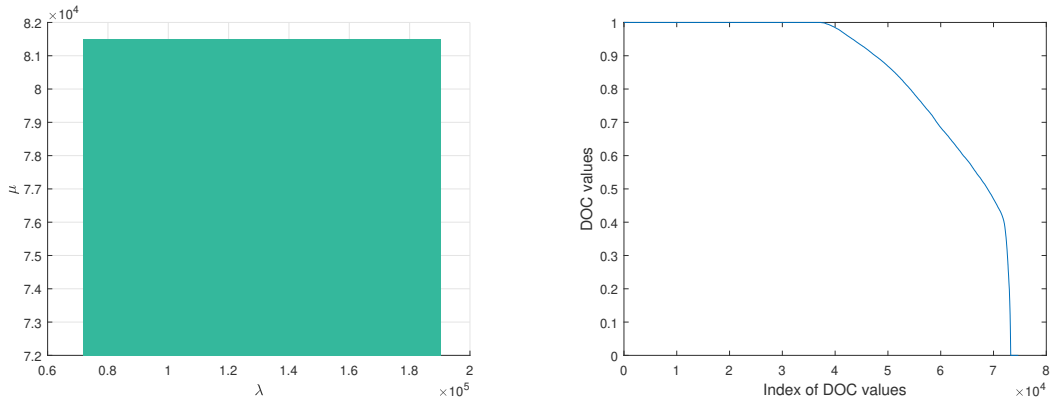


Figure 3.5 – Measurement selections with $GDOC$ method



(a) (b)

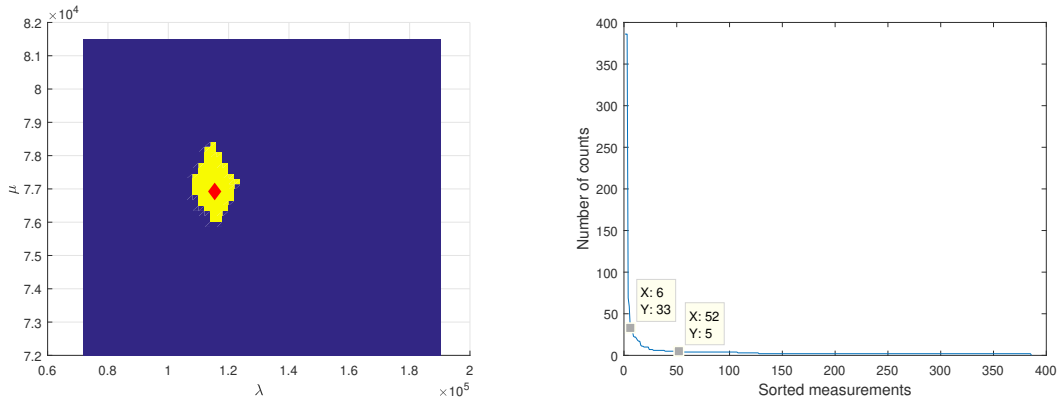
Figure 3.6 – Identification with $Count$ method with noise in the data. (a) Empty solution set. (b) DOC value of all possible pair of measurements.

Figure 3.4(a) then shows that in the case of noisy data, the identified set (green color) is empty due to inconsistency within the measurements, hence the need to remove some measurements. Figure 3.5 shows the function f_{GDOC} (see Equation (2.15)), providing the $GDOC$ values of all the measurements in decreasing order. It can be observed that the value of $GDOC$ starts to decrease abruptly on this example where $GDOC$ is between 0.65 and 0.52. As indicated in Algorithm 2, a possible criterion to select measurements could be to detect this abrupt decrease of the estimator of consistency of the data. This detection should of course be performed automatically, and will be so in our statistical analysis, but in this illustrative case the threshold was manually set at 0.65. The corresponding identified set (yellow color) is presented in Figure 3.4 (b), where 86 measurements were removed. The solution set in yellow still contains the reference values $\lambda_0 = 1.15 \cdot 10^5 MPa$ and $\mu_0 = 7.69 \cdot 10^4 MPa$ shown by red mark, albeit it is close to the border of the set.

3.3.1.2 Count method

When all measurements are consistent and noiseless, the count method provides just the same solution as the first approach. We therefore only show the results in the noisy case.

In this case, Figure 3.6(a) shows that the identified set (green color) is empty due to inconsistency within the measurements. It is also obvious from Figure 3.6(b) that displays ordered values of the DOC consistency measure (Equation (2.10)) for all pairs of measurements. Indeed, a lot of these values are very low, including some that have value zero (meaning that some pairs of measurements are completely inconsistent with one another).



(a) (b)

Figure 3.7 – Identification with $Count$ method and $Count$ curve. (a) Feasible set of parameters. (b) $Count$ curve

We have therefore applied Algorithm 3 to this case, using a DOC threshold τ of 0.1, by observing DOC curve in Figure 3.6(b). Figure 3.7(b) shows the counts of all the measurements in decreasing order, i.e., the function f_{Count} given by Equation (2.17), and it can be observed that the value of count starts to decrease abruptly on this example when the count value is between 33 and 5. Again, identifying this inflection will be performed automatically in our analysis, yet in this first example, the threshold was set to 5 counts. The corresponding identified set (yellow color) is presented in Figure 3.7(a), where 46 measurements were removed whose count is strictly greater than 5. The solution set in yellow still contains the reference values $\lambda_0 = 1.15 \cdot 10^5 MPa$ and $\mu_0 = 7.69 \cdot 10^4 MPa$ shown by red mark, and is smaller than the one obtained with the first method.

3.3.1.3 $GCONS$ method

Finally, Figure 3.8 shows the solution set resulting from applying Algorithm 5 based on the $GCONS$ function with a threshold $\tau = 0.1$. Again, it still contains the reference values $\lambda_0 = 1.15 \cdot 10^5 MPa$ and $\mu_0 = 7.69 \cdot 10^4 MPa$, that are on the border of the set, despite it being significantly larger. The feasible set (yellow color) of the identified parameter using $GCONS$ method was obtained by removing 55 measurements.

We saw that the $GDOC$, $Count$ and $GCONS$ methods remove 86, 46, and 55 measurements, respectively, but a question that remains is whether the methods remove the same measurements, or really differ qualitatively? To measure the similarity in terms of which measurements are removed by any two methods, we use the Jaccard index (J), also known as Intersection over Union, for two sets. We denote the three sets of removed measurements for $GDOC$, $Count$, and

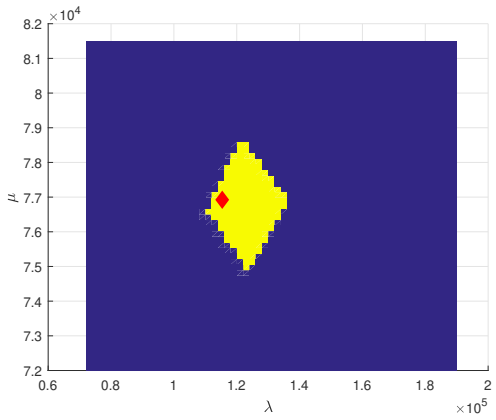


Figure 3.8 – Feasible set of parameters ($GCONS$)

$GCONS$ method as $\mathbb{R}_1$, $\mathbb{R}_2$, and $\mathbb{R}_3$, respectively. We found that $\mid \mathbb{R}_1 \cap \mathbb{R}_2 \mid = 40$, $\mid \mathbb{R}_1 \cup \mathbb{R}_2 \mid = 92$, $\mid \mathbb{R}_1 \cap \mathbb{R}_3 \mid = 49$, $\mid \mathbb{R}_1 \cup \mathbb{R}_3 \mid = 92$, $\mid \mathbb{R}_2 \cap \mathbb{R}_3 \mid = 31$, $\mid \mathbb{R}_2 \cup \mathbb{R}_3 \mid = 70$.

Table 3.1 – Jaccard Index(J) to measure the similarity in terms of which measurements are removed by any two methods

Methods	Jaccard Index(J)
$GDOC$ and $Count$	$J(\mathbb{R}_1, \mathbb{R}_2) = \frac{\lvert \mathbb{R}_1 \cap \mathbb{R}_2 \rvert}{\lvert \mathbb{R}_1 \cup \mathbb{R}_2 \rvert} = 0.4348$
$GDOC$ and $GCONS$	$J(\mathbb{R}_1, \mathbb{R}_3) = \frac{\lvert \mathbb{R}_1 \cap \mathbb{R}_3 \rvert}{\lvert \mathbb{R}_1 \cup \mathbb{R}_3 \rvert} = 0.5326$
$Count$ and $GCONS$	$J(\mathbb{R}_2, \mathbb{R}_3) = \frac{\lvert \mathbb{R}_2 \cap \mathbb{R}_3 \rvert}{\lvert \mathbb{R}_2 \cup \mathbb{R}_3 \rvert} = 0.44297$

From the results reported in Table 3.1, we can see that the Jaccard index values are similar for all pairs of method, and also relatively high. This shows that the methods will, in general, remove different measurements (with of course some overlap), which means that comparing their efficiencies in different settings, as done in the next sections, makes sense.

In the next section, we perform a statistical comparison of these different approaches, comparing them to the Bayesian approach, as well as in term of precision and accuracy.

3.3.2 Comparison of selection methods with noisy measurements

We will now make a statistical comparison of our three selection methods on the same numerical example, as well as with a standard Bayesian approach, in the case where measurements are noisy.

We will apply the set-valued inverse method and Bayesian inference method to identify the elastic properties (Lamé parameters: λ and μ) of a homogeneous

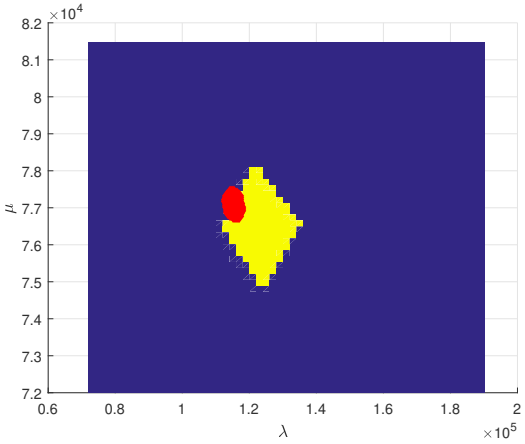


Figure 3.9 – Feasible sets of parameters ($GDOC$ + Bayes)

2D plate with 386 measurements. For the set-valued inverse method, information on the measurement $\tilde{y}$ is described in an interval form: $[\ \tilde{y} - 2\sigma,\quad \tilde{y} + 2\sigma\ ]$ with $\sigma = 0.0020$ and prior information about the parameters ($\mathbb{S}_\theta^0$) is considered as a $2D$ box $\lambda^p \times \mu^p$ with $\lambda^p = [0.72 \cdot 10^5, 1.90 \cdot 10^5]MPa$ and $\mu^p = [\ 7.2 \cdot 10^4, 8.15 \cdot 10^4]MPa$. For the Bayesian inference method, error on the measurement $\tilde{y}$ is modeled by a Gaussian white noise:$\sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 0.0020$ and the prior information about the parameter is modeled by a uniform distribution:$U_\lambda(0.72 \cdot 10^5, 1.90 \cdot 10^5)MPa$, $U_\mu(7.2 \cdot 10^4, 8.15 \cdot 10^4)MPa$. This means that the Bayesian model is in this case well specified, and should perform well.

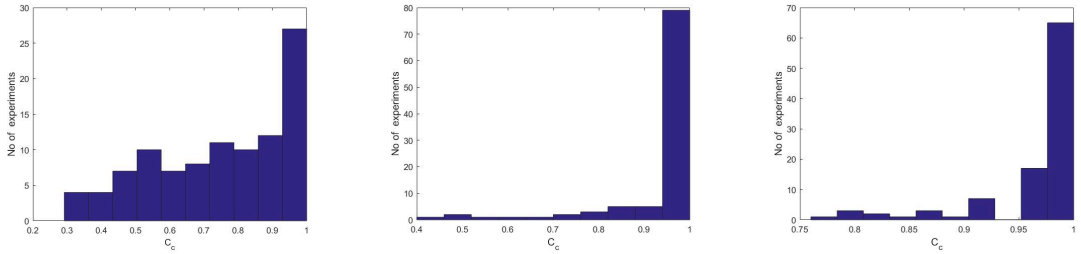
3.3.2.1 Comparison with Bayesian inference

To perform a comparison between our set-valued approach and a Bayesian inference method, we will transform the resulting posterior distribution on parameters of the Bayesian inference in a set-valued estimate. To do so, we will estimate a 90% credibility set of the Bayesian approach.

Figure 3.9 shows the feasible set (yellow color) of the identified parameter using $GDOC$ method with 86 measurements removed and the feasible set (red color) corresponding to the 90% credibility set of the Bayesian approach. The results showed that both identified sets are not completely consistent with one another and we can define the degree of consistency (C_c) between them as

$$C_c = \frac{\mathcal{A}(S_s \cap B_s)}{min(\mathcal{A}(S_s), \mathcal{A}(B_s))} \tag{3.17}$$

where S_s and B_s represent the solution set corresponding to the set-valued inverse method and the Bayesian inference method, respectively. This comes down to apply Equation (2.10) to the two different approaches.



a) $GDOC$ method b) $Count$ method c) $GCONS$ method

Figure 3.10 – Comparison of selection methods in terms of consistency(C_c) with Bayesian inference method

To make our comparison, we generated 100 experiments with 386 measurements to which were added a Gaussian white noise:$\sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 0.0020$ for each experiment, that is

$$\tilde{y} = y_{ref} + \mathcal{N}(0, \sigma^2).$$

The same 100 experiments have been used while comparing all three methods with respect to Bayesian inference method. In the case of $GDOC$ and $Count$ method , we use Matlab function ($knee\_pt.m$ [Kaplan, 2021]) to detect the bend of the curve and automatically detect measurements to remove while performing the 100 experiments. In the case of $GCONS$ method, we set the threshold τ used in Algorithm 5 to 0.1.

Table 3.2 – Comparison of selection methods in terms of consistency (C_c) with Bayesian inference method

Method	Mean of C_c	Standard deviation of C_c
$GDOC$	0.75	0.20
$Count$	0.94	0.12
$GCONS$	0.96	0.05

Figures 3.10(a) to 3.10(c) show the distribution of consistency values C_c for the three methods. The results show that all three methods are in general in good agreement with the Bayesian inference approach when this latter makes the right model assumption. Yet one easily see that the $GDOC$ distribution is much less peaked, and has much more elements far away from one. This is also visible in the mean and variance of the distributions reported in Table 3.2: the mean is much lower and the variance much higher for Algorithm 2. We suspect that this is due to the fact that the algorithm does not really account for interactions between removed measurements.

The fact that the C_c value is quite high for the $GCONS$ and $Count$ methods

Table 3.3 – Comparison of selection methods in terms of DOI between set-valued inverse method and Bayesian inference method

Method	Mean of $DOI_{B_s \subseteq S_s}$	Standard deviation of $DOI_{B_s \subseteq S_s}$	Mean of $DOI_{S_s \subseteq B_s}$	Standard deviation of $DOI_{S_s \subseteq B_s}$
$GDOC$	0.75	0.20	0.19	0.0445
$Count$	0.94	0.12	0.32	0.13
$GCONS$	0.96	0.05	0.16	0.0150

indicate that they are often consistent with the 90% credible set of the Bayesian approach, hence that one is often included in the others. Figure 3.9 suggests that in practice, the Bayesian approach is often more precise than our proposed methods. This is confirmed by the numbers in Table 3.3, that reports the degrees of inclusion(see Equation 2.10) of the solution set $B_s \subseteq S_s$ and vice-versa for the 100 experiments. We can easily see that the first two columns are the same as the numbers in Table 3.2, while the numbers for $DOI_{S_s \subseteq B_s}$ are much lower. This confirms that, in practice, the Bayesian approach will often induce a more precise solution set, which could be expected by the fact that our approaches reject measurements, hence rely on less information, while the Bayesian approach tries to build a compromise between all measurements.

3.3.2.2 Precision and accuracy of selection methods

Let us now compare the methods between themselves, first in term of precision or size of the obtained set, and then in terms of accuracy. Figure 3.11 shows the distribution of the sizes of the obtained solution sets for the three approaches, and Table 3.4 provides summary statistics (mean and variance) of these distributions. From this, we can conclude that while $GCONS$ method provides sets with a more stable size, such sets also tend to be larger than the ones obtained with the two other approaches, with the $Count$ method tending to give the smallest sets. We can explain the observed variance of the $GDOC$ and $Count$ methods by the fact that the inflection point, used as a threshold, is much more likely to vary from experiment to experiment.

Table 3.4 – Comparison of selection methods: statistical summary

Method	Mean of Size of set	Standard deviation of Size of set
$GDOC$	95.18	33.65
$Count$	79.56	35.87
$GCONS$	142.54	7.08

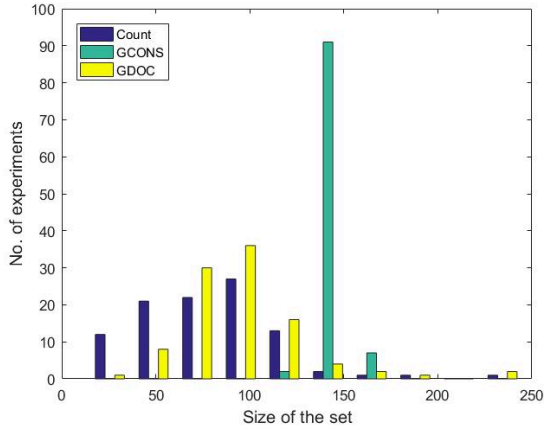


Figure 3.11 – Comparison of selection methods in terms of the size (number of grid points) of the solution set

A second aspect that is important in identification methods is whether the approaches can retrieve the true solution. For this, we computed the accuracy value

$$A_C = \frac{|\{Sol^i : \theta^* \in Sol^i\}|}{K} \tag{3.18}$$

where Sol^i is the solution set obtained for the ith experiment, and K is the total number of experiments. Table 3.5 provides such an information, as it summarises the number of experiments for which the true parameter values are included in the solutions sets. Both the $Count$ and $GCONS$ methods retrieve the true solution in all experiments, while for the $GDOC$ method, this is only true for 90 experiments, which is comparable to what the Bayesian inference would achieve with more precise solution sets, as we take the 90% credibility intervals. So in this case both the $Count$ and $GCONS$ methods clearly provide superior results, at least in terms of accuracy. Coupled with the fact that the $Count$ method provides more precise sets in general, this method clearly seems preferable.

Table 3.5 – Comparison of selection methods in terms of accuracy

Method	A_C
$GDOC$	90
$Count$	100
$GCONS$	100

Figure 3.12 provides a complementary information in the form of the distributions of number of measurements that are removed from the initial set to obtain the final solutions. It clearly shows that the $GCONS$ and $Count$ methods tend to remove fewer measurements than the $GDOC$ approach. Together with the results

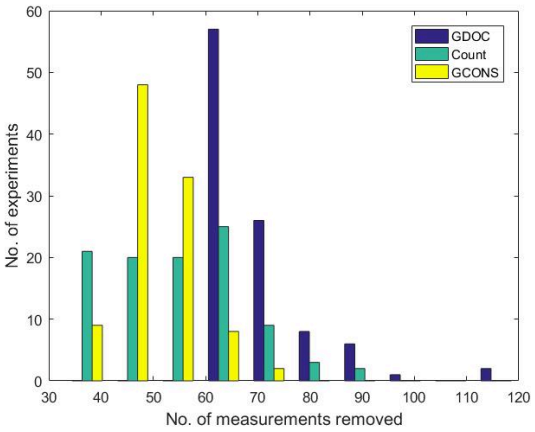


Figure 3.12 – Comparison of selection methods in terms of the number of measurements removed

presented in Figure 3.11, this indicates that the number of removed measurements is not directly linked to the size of the obtained final sets, as $GCONS$ was giving the larger sets, while generally removing less measurements than $GDOC$.

To summarise, it appears that the three methods provide results that are qualitatively comparable, but remain quantitatively different. Accounting for interactions seem to increase results reliability and accuracy, while choosing a data- or experiment-dependent selection threshold seem to result in smaller solution sets. The price to pay for a high reliability, compared to a well-specified Bayesian approach, is a higher conservativeness. So, in those cases where the Bayesian model is well-specified and we do not require strong guarantees, Bayesian inference seems preferable. Let us now look at what happens when the Bayesian model is misspecified, i.e., do not perfectly represent data and measurement processes.

3.3.3 Comparison in presence of outliers and misspecified Bayesian model

In this section, we compare the set-valued inverse methods and the Bayesian inference method in presence of outliers i.e, when some data become aberrant. To do this, we use 8 sets of 100 experiments (each experiment with 386 measurements) in a way such that for each set the percentage of outlier measurements will increase. In practice, we use the following scheme:

$$\tilde{y}_0 = y^{Ref} + \varepsilon \tag{3.19}$$

$$\tilde{y} = \tilde{y}_0 + \alpha I \varepsilon \tag{3.20}$$

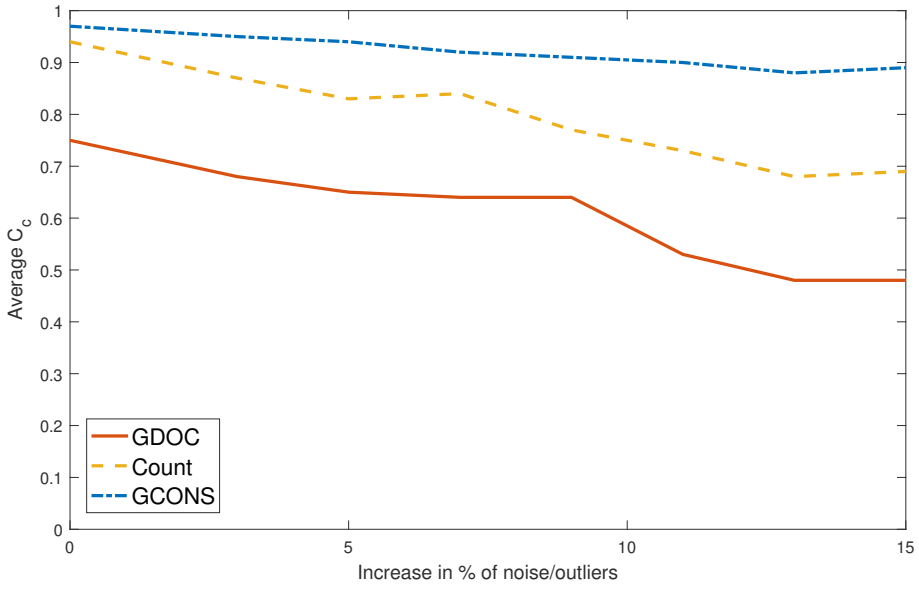


Figure 3.13 – Sensitivity to outliers

where $\tilde{y}_0$ are noisy measurements similar to the one of the previous section, with $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ being the initial noise. We then create outliers by considering $\alpha \in \mathbb{R}^+$ as a multiplicative factor of the initial noise ε, applied to an average proportion p_i of the initial noisy measurements, this average proportion being obtained by setting $I \sim \mathcal{B}(p_i)$ as a Bernoulli variable with parameter p_i that will vary between sets of 100 experiments. For our experiments, we set $\alpha = 5$ and we used the values 0%, 3%, 5%, 7%, 9%, 11%, 13%, 15% for p_i, starting from no outliers to an average of 15%.

For each set of experiments (thus for 800 experiments), we have performed identification using our set-valued inverse methods and Bayesian inference. Figure 3.13 reports the average values of the consistency coefficient C_c given by Equation (3.17) for all 100 experiments and each percentage of outliers. From this, one can easily see that for all our three methods, the consistency with the Bayesian approach decreases as the percentage of outliers increases. This is less severe for the $GCONS$ approach, as this latter one tends to provide larger solution sets, that more often include the Bayesian solution. Clearly, these graphs show that in presence of outliers, the probabilistic approach and our methods behave differently.

However, if the probabilistic and set-valued methods provide different results, it remains to determine which one gives the best. Figure 3.14 provides some answer to that, by reporting the average number of times that each method includes the true parameter values, given by the value A_c. It can be observed that as the

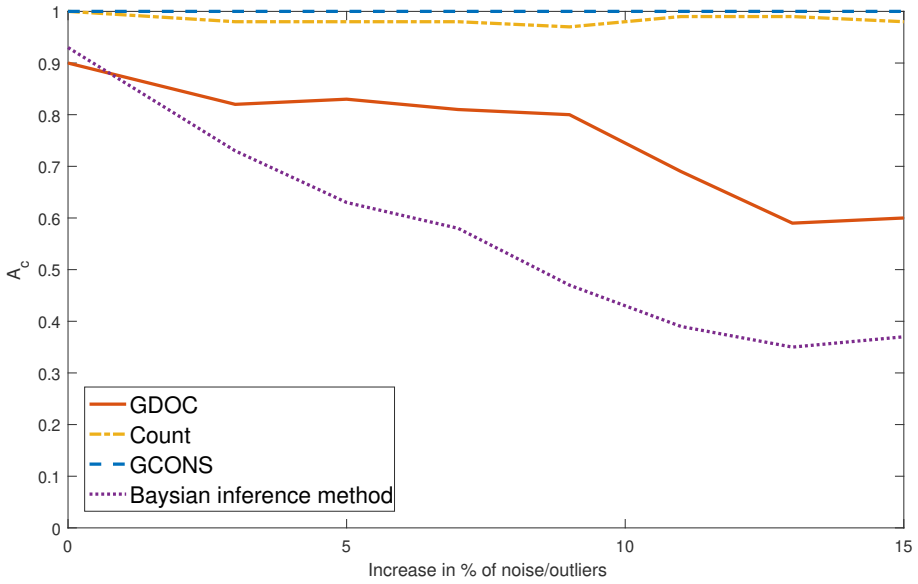


Figure 3.14 – Consistency with exact parameter values

quantity of outliers increase, the performances of both the Bayesian and the $GDOC$ approaches decrease dramatically, to the point that the 90% credible Bayesian region contains the true parameter only about 50% of the time, and this with only 15% of outliers. $GDOC$ is also sensitive to outliers, albeit to a lesser extent. Again, this is probably due to the fact that $GDOC$ does not really consider interactions between measurements when removing them, something that seems essential to detect outliers. In contrast, both the $Count$ and $GCONS$ methods behave remarkably well, as the $Count$ method nearly always contains the true parameter values, while $GCONS$ always contains them.

Of course, one could include the outliers characteristic in a Bayesian approach and obtain a well-specified model, but this would mean being able to assess the nature and quantity of such outliers, as well as a quite higher computational burden. In contrast, the $Count$ and $GCONS$ approaches can deal with them with an identical computational price, and without any knowledge of the nature and quantity of outliers.

This experiment shows that the two methods intrinsically differ: while the Bayesian approach requires reliable probabilistic information in order to work and will consider every measurement as informative, our approach rely on less assumptions (only requiring boundary values) and can consider some measurements are completely uninformative or even damaging for the inference. The Bayesian approach requiring stronger assumptions and knowledge, it will usually perform better (i.e., provide more precise and equally accurate results) in those cases where

these assumptions are valid, but will also possibly suffer an important bias in case they are not. Our method seems therefore advisable in those sensitive engineering or mechanical applications where this knowledge is not present, so as to avoid possible bias.

We also check how effectively the three outlier detection methods are removing some of the noisy measurements which are converted into outliers. For a given experiment, we denote the list of such outliers by $\mathbb{L}_{p_i}$, and by $\mathbb{R}_e$ the list of total inconsistent measurements we remove by each method. For each experiment, we compute the two values

$$C_e = \frac{\mid \mathbb{L}_{p_i} \cap \mathbb{R}_e \mid}{\mid \mathbb{L}_{p_i} \mid},$$

and

$$C_r = \frac{\mid \mathbb{L}_{p_i} \cap \mathbb{R}_e \mid}{\mid \mathbb{R}_e \mid}$$

, that respectively compute the percentage of outliers we get within the removed measurements, out of the total number of possible outliers, and the percentage of those removed measurements that are outliers. Before commenting the results, detailed in Tables 3.6 and 3.7, two remarks are in order: first, given the way we generate outliers, it is clear that an identified outlier for which the value $I = 1$ in Equation (3.20) may actually suffer from a very small noise (in case the random noise ε is small), so may not truly be an outlier; second, given that non-outliers also suffer from a random normal, possible unbounded noise, those latter will also induce inconsistencies (potentially even more important than designed outliers).

Table 3.6 – The effectiveness of three outlier detection methods to remove some of the noisy measurements which are converted into outliers.

Method	Average value of C_e over 100 experiments							
-	0%	3%	5%	7%	9%	11%	13%	15%
$GDOC$	–	0.8350	0.8338	0.8267	0.8143	0.7979	0.79	0.7894
$Count$	–	0.8040	0.8015	0.7967	0.7864	0.7867	0.7849	0.7835
$GCONS$	–	0.5070	0.50	0.5033	0.5018	0.5048	0.4996	0.5115

However, even accounting for these two facts, we can see in Table 3.6 that all methods are able to capture a fair number of the designed outliers (i.e., those measurements for which $I = 1$), and that this number is stable for different percentage of outliers. From the results, it seems that the $GCONS$ approach tends to select less outliers, maybe because the underlying measure does not rely on pairwise inconsistency measures, aggregated or not. Indeed, one may expect outliers to be strongly detected by pairwise measure, as by definition they

are inconsistent with most single measurements. It is also not surprising to not reach 100% detection, given the general remarks we have made about the outlier simulation process.

Table 3.7 – The $\%$ of outliers among the removed measurements in $\mathbb{R}_e$

Method	Average value of C_r over 100 experiments							
-	0%	3%	5%	7%	9%	11%	13%	15%
$GDOC$	–	0.1169	0.1474	0.1917	0.2749	0.4114	0.4711	0.4836
$Count$	–	0.1650	0.2099	0.27	0.3636	0.4633	0.5029	0.5138
$GCONS$	–	0.0949	0.1203	0.1619	0.2347	0.3240	0.3626	0.3829

Results of Table 3.7 are in line with expectation, as the percentage of outliers within the removed measurements increases as the number of outliers increases as well. Indeed, we cannot expect C_r to be high when having 3% outliers, as 3% of 386 means about 10 measurements, while the methods tends to remove about 50 to 100 measurements (in this sense, the numbers in the 3% column of Table 3.7 are in line with expectation). The increase in the percentages show that all methods are able to catch more outliers as their quantity grows.

3.3.4 Choice of the threshold value τ in the selection methods

We discussed the three selection methods in the previous Section to remove inconsistent or outlier measurements. The common factor in all three methods is that we need to choose some threshold value τ to use these methods. All the definitions of consistency measure we discussed in Section 2.2.2 tell us in a general sense to choose a low threshold value. We can set a general criterion to choose a data or experiment-dependent threshold value. For example, in the $GDOC$ and $Count$ method, we can always rely on the $GDOC$ and DOC curve to get an idea about the threshold value. But, in the case of the $GCONS$ method, we do not have any predefined idea about choosing a threshold value except it should be a low value. We propose that we can use the DOC curve to choose the threshold value for the $GCONS$ method. $GCONS$ is a global measure, but it turns into a pairwise measure in the case of two measurements (see 2.3.2.2, property 4). This gives us insight that the DOC curve can help choosing a threshold value for the $GCONS$ method.

We will now illustrate why it is advisable to choose a low DOC threshold value for the $Count$ method. We considered the same example as in Section 3.3.1.2. Figure 3.15 shows the ordered values of DOC consistency measure (Equation 2.10) for all pairs of measurements. The Figure 3.16 (a,b,c,d,e,f,g) shows the $Count$

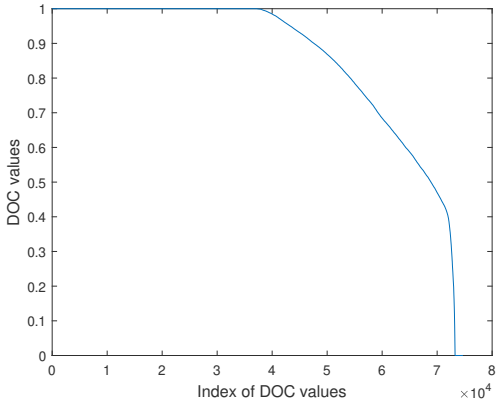


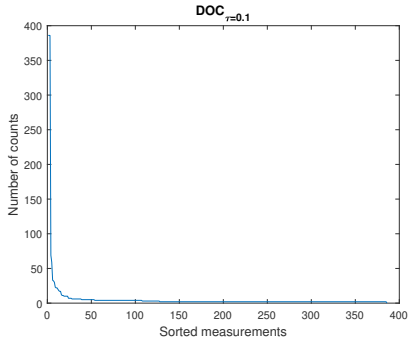
Figure 3.15 – DOC value of all possible pair of measurements(DOC curve to choose threshold τ.

curve for the different threshold values, such as τ = 0.10,0.3,0.4,0.52,0.81 and 0.96. We have observed that as the threshold value of DOC increases $Count$ curve starts to become unstable i.e., in terms of smoothness of the curve. The Count curve in Figure 3.16(a) has a smooth nature in contrast to Figure 3.16(g), and it is complicated to select the number of inconsistent measurements based on the count value in Figure 3.16(g). Hence, to effectively obtain the count of the measurements, i.e. $Count$ curve, one should use a low threshold value of DOC, the reason is that the low value of DOC is a kind of indicator of low consistency between the two measurements. It is possible to decide the low threshold value of DOC by observing the DOC curve's nature shown in Figure 3.15.

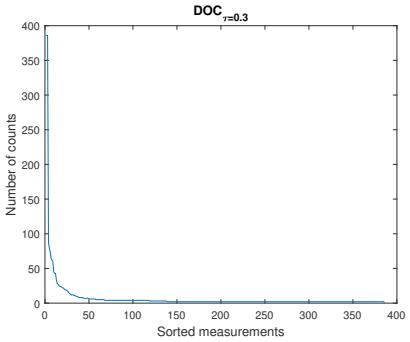
3.4 Summary

In this chapter, we have considered the issue of solving the inverse problem in the presence of interval-valued uncertainty in the measurements. While most common strategies such as least-square minimization or Bayesian inference intend to find a solution that is not too far away from all the observed measurements, we considered the possibility of removing some of the measurements, only keeping a subset of measurements for which the inverse problem solutions were consistent. While such a view is not entirely new and is for example at work in the Q-intersection method [Drevelle and Bonnifait, 2012], this latter solution is not really applicable when the number of measurements is high, which is quite often the case in mechanical problems.

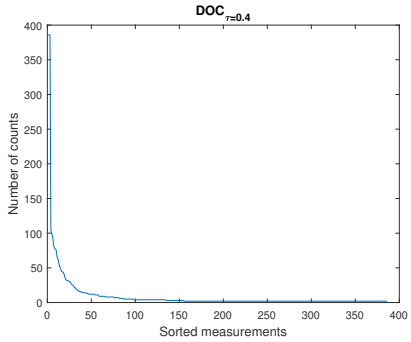
The selection (outlier detection) methods we applied to detect inconsistent measurements have the advantage that they are easy to compute, making them affordable even for a large number of measurements.



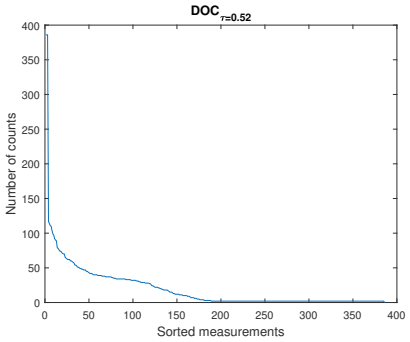
a) Count curve with $DOC_{\tau=0.1}$



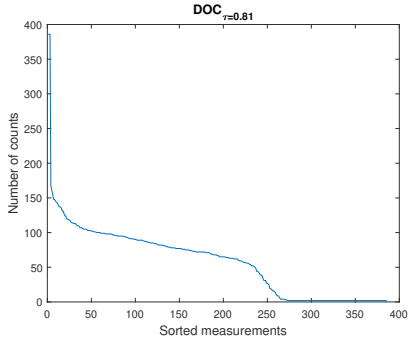
b) Count curve with $DOC_{\tau=0.3}$



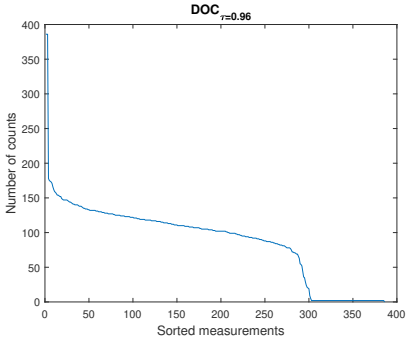
c) Count curve with $DOC_{\tau=0.4}$



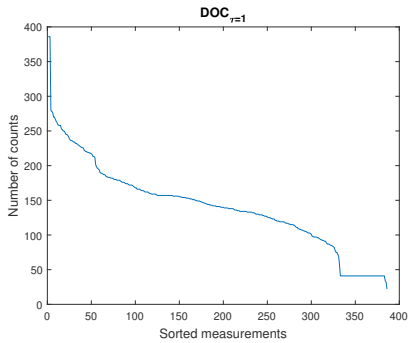
d) Count curve with $DOC_{\tau=0.52}$



e) Count curve with $DOC_{\tau=0.81}$



f) Count curve with $DOC_{\tau=0.96}$



g) Count curve with $DOC_{\tau=1}$

Figure 3.16 – Count curve for the different threshold values

We applied our various strategies to identify the elastic properties of homogeneous isotropic material, comparing them to a Bayesian model in two situations: one where the Bayesian model was accurate, the other where it was misspecified. Our main conclusions from these experiments are that:

- The Bayesian and set-valued approaches deliver solutions of similar quality when the former makes the right assumptions, but the Bayesian approach is quite sensitive to misspecification and the presence of outliers, while our approaches are much more robusts.
- When selecting measurements, it seems essential to integrate possible interactions between the removed measurements to ensure a good identification, especially when the presence of outliers of anomalies is suspected.

In the next chapter, we will extend the application of the set-valued inverse method to identify parameters when measurements are in a large amount and damage detection in a material.

Application of set-valued inverse method: in a higher dimensional setting and to detect damage in the material.

Contents

4.1 Introduction

This chapter presents two applications of the set-valued inverse method:

- Identification from full-field displacement measurements when they are in a large amount.
- Damage detection of the material

Any identification strategy becomes computationally complex when we have either a large number of parameters to identify or of available measurements. To tackle this issue, we illustrate the use of a non-uniform(sparse) grid of points to describe the prior information about parameters and surrogate modeling. In the last chapter, our focus was on removing inconsistent or outlier measurements to obtain the solution to the inverse problem as the inconsistency was only due to measurement error. However, as we shall see here, our outlier detection methods, can help to distinguish between outliers due to measurement and model error. In the damage detection problem, we solve the inverse problem and detect outlier measurements due to model error that indirectly determines the location of damage in the material.

4.2 Identification in a higher dimensional context

The mechanical inverse problem that we considered in Chapter 3 consists of identifying the two parameters with a small number of measurements. This problem can become computationally complex in three cases:

1. a larger number of parameters to identify,
2. with the large FE model,
3. large data set or available measurements are numerous.

In this section, we will discuss the methods to improve computational efficiency of the set-valued inverse method in a higher dimensional context such as

1. sparse grid of points using Latin hypercube sampling (LHS),
2. surrogate model,

respectively with targeted applications:

- higher dimensional parameter space,
- large FE model.

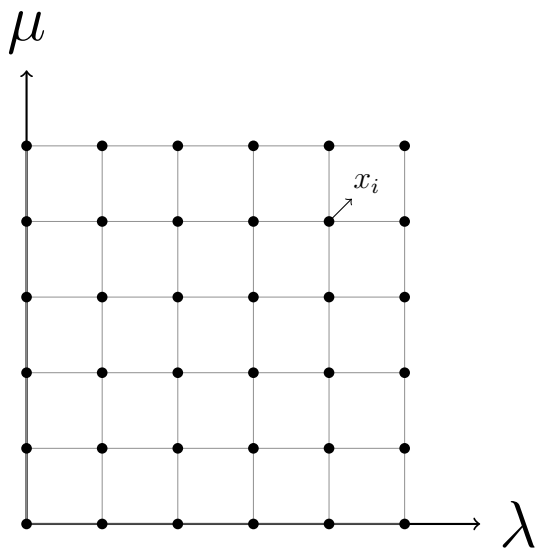


Figure 4.1 – Description of parameters with a grid of points.

4.2.1 Higher dimensional parameter identification

In the case of 2D parameter identification and with a small number of measurements, we described the prior information about parameters with a uniform grid of points. Figure 4.1 shows the 2D box of the grid of points, where the coordinates represents the values of the parameters λ, μ, and x represents the position of the grid point. The description shown is uniform(regular), but it can be non-uniform. When considering higher dimension, for the uniform description of the grid of points, the number of grid points are given as

$$N_g = n_g^M \tag{4.1}$$

where N_g is the number of grid points, n_g is the number of chosen points for each dimension, M is the number of parameters.

In our application, we need to solve the Finite Element(FE) model at each grid point to compute displacements corresponding to the parameters. If the inverse problem consists of identification of a large number of parameters, then that leads to high computational cost. One way to reduce the computational cost, in the case of higher dimensional parameter identification, is to describe prior information about parameters with a non-uniform grid of points. There are various methods to generate the grid of points non-uniformly, such as Halton sequences [Halton, 1960, Wang and Hickernell, 2000] or Latin hypercube sampling (LHS) [McKay et al., 1979, Saliby and Pacheco, 2003, Shields and Zhang, 2016]. [Sui, 2017] compares how accurately one can describe the set with a non-uniform grid of

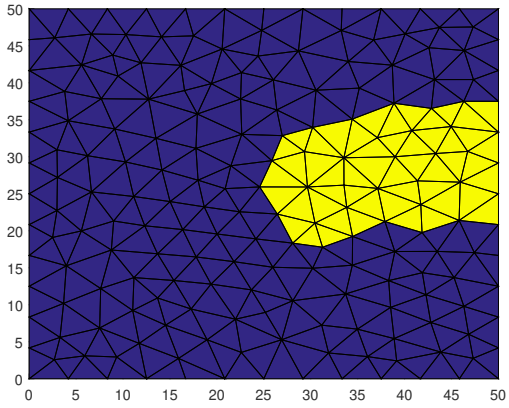


Figure 4.2 – Heterogeneous material

points using LHS and Halton sequences compared to the regular(uniform) grid of points. In this work, we use the LHS sampling method to describe prior information about parameters with the non-uniform description of the grid of points. Such a non-uniform description of a grid of points reduces the number of times one has to solve FE calculations because the number of FE calculation is equal to the number of grid points. Let us considere an example of the description of four parameters with uniform description, with $n_g = 60$ and $M = 4$, the number of grid points, N_g, is equal to 12960000 due to the so-called curse of dimensionality.

Now, we illustrate how we can solve 4D identification problem with the use of non-uniform description of the grid of points. We considered the application where we want to identify 4 parameters (λ_1, μ_1, λ_2, μ_2) of a heterogeneous material shown in Figure 4.2. λ_1 and μ_1 represents the properties of the material described with blue color and λ_2 and μ_2 represents the properties of the material described with yellow color.

To perform the identification using the set-valued inverse method of a heterogeneous 2D plate under plane strain as shown in Figure 4.2, we simulate a plate clamped on the left side and loaded on the right side by a uniform traction $f = 1000\ N/m$. To generate displacement measurement data $\tilde{y}$ (386 measurements), exact displacement data y^{Ref} are simulated by a Finite Element(FE) model(193 nodes, 336 elements) as shown in Figure 4.2 considering the reference values $\lambda_{01} = 1.15 \cdot 10^5 MPa$, $\mu_{01} = 7.69 \cdot 10^4 MPa$ and $\lambda_{02} = 5.76 \cdot 10^4 MPa$, $\mu_{02} = 3.84 \cdot 10^4 MPa$. We also consider a Gaussian white noise with 0 mean (no systematic bias) and with standard deviation σ, taken here as 5% of the average of all the exact displacement values. To apply our methods, the uncertainty on the measurements has to be given in interval form. Therefore, each measurement is modelled as $[\tilde{y}_k - 2\sigma, \tilde{y}_k + 2\sigma]$. The width of 2σ ensures that sufficient measurements

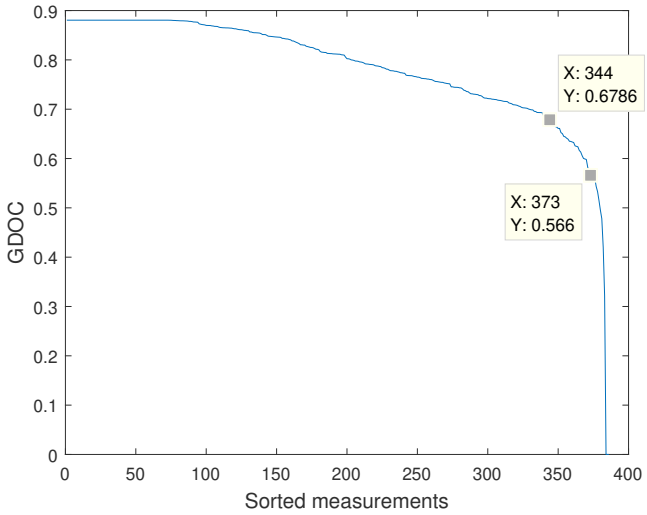


Figure 4.3 – Measurement selections with $GDOC$ method (4D identification)

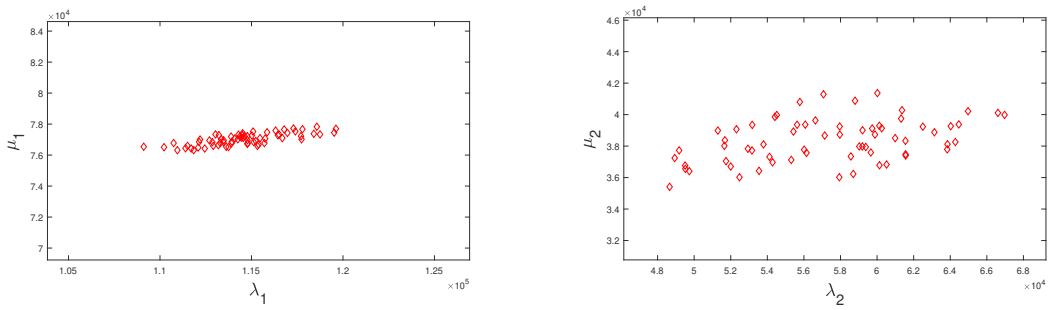


Figure 4.4 – 4D identification with $GDOC$ method. (a) Solution set($\lambda_1 \times \mu_1$). (b) Solution set ($\lambda_2 \times \mu_2$).

will be consistent with one another. Prior information about the parameters ($\mathbb{S}_\theta^0$) is considered as a $4D$ box described by a grid of points $\lambda_1^p \times \mu_1^p \times \lambda_2^p \times \mu_2^p$ with $\lambda_1^p = [1.03 \cdot 10^5, 1.26 \cdot 10^5]MPa$, $\lambda_2^p = [4.61 \cdot 10^4, 6.92 \cdot 10^4]MPa$ and $\mu_1^p = [\,6.92 \cdot 10^4, 8.46 \cdot 10^4]MPa$, $\mu_2^p = [\,3.07 \cdot 10^4, 4.61 \cdot 10^4]MPa$. With LHS sampling method, to describe $4D$ box of non-uniform grid of points, we used 20000 points for the grid.

We solved the inverse problem with the random noise in the data. Figure 4.3 shows the function f_{GDOC} (see Equation (2.15)), providing the $GDOC$ values of all the measurements in decreasing order. It can be observed that the value of $GDOC$ starts to decrease abruptly on this example when $GDOC$ is between 0.67 and 0.56. As indicated in Algorithm 2, a possible criterion to select measurements could be to detect this abrupt decrease of the estimator of consistency of the data. The threshold was manually set at 0.65. We obtain the corresponding identified set of 4 parameters such that the solution set($\mathbb{S}_\theta$)=$\lambda_1 \times \mu_1 \times \lambda_2 \times \mu_2$. Figure 4.4 represents the solution set in terms of the projection of $\mathbb{S}_\theta$ on λ_1, μ_1, and λ_2, μ_2

axis respectively, where 34 measurements were removed. The solution set still contains the reference values $\lambda_{01} = 1.15 \cdot 10^5 MPa$, $\mu_{01} = 7.69 \cdot 10^4 MPa$ and $\lambda_{02} = 5.76 \cdot 10^4 MPa$, $\mu_{02} = 3.84 \cdot 10^4 MPa$. We solved 4D identification using a non-uniform description of the grid of points. We used a 20000 number of grid points, which is quite reasonable if we compare it to 12960000 points in the case uniform description when M=4. In the above example, with 20000 points, we get a reasonable solution set. There is always a question of how many points one should choose to describe a higher dimension set accurately with a non-uniform grid of points. We did not do a detailed study on this topic and chosen a sufficient number of points such that we will not get an empty solution set in the context of the set-valued inverse method. [Sui, 2017] has proposed a method to choose the number of points to describe the set accurately.

4.2.2 Identification with a large FE model

In practice, the mechanical inverse problem based on DIC data consists of solving an inverse problem with thousands of measurements, which is quite higher than the number we considered in the application of chapter 3, where we used a minimal number of measurements to illustrate different outlier detection criteria. In this section, we identify the elastic properties (Lamé parameters: λ and μ) of a homogeneous 2D plate under plane strain, as shown in Figure 3.2 with a large FE model. In our numerical applications, a large FE model(mesh) with a large number of nodes and elements means that a large number of measurements (number of measurements = number of nodes $\times$ 2). But in practice, it is not always the case; in DIC measurements, measurement data grids and FE mesh are not the same.

Figure 4.1 shows the 2D box of the grid of points, where each coordinates represents the values of the parameters λ, μ, and x represents the position of the grid points. In our application, we need to solve the Finite Element(FE) model at each grid point to compute displacements corresponding to the parameters. If the inverse problem consists of identification from a large FE model with a large number of nodes and elements, then that leads to high computational cost. To reduce the computational cost, i.e., to avoid solving FE calculations at each grid point, we use the radial basis function (RBF) interpolation method [Ozcanan and Atahan, 2020, Toja-Silva et al., 2014, Bajer and Holena, 2012] to build a surrogate model. In this method, we solve the FE calculations at a few sample grid points, and then using the surrogate model, we compute an approximated displacement at any grid point that helps to reduce computational cost. Now we describe the steps to build the surrogate model to compute the displacement at any grid point using

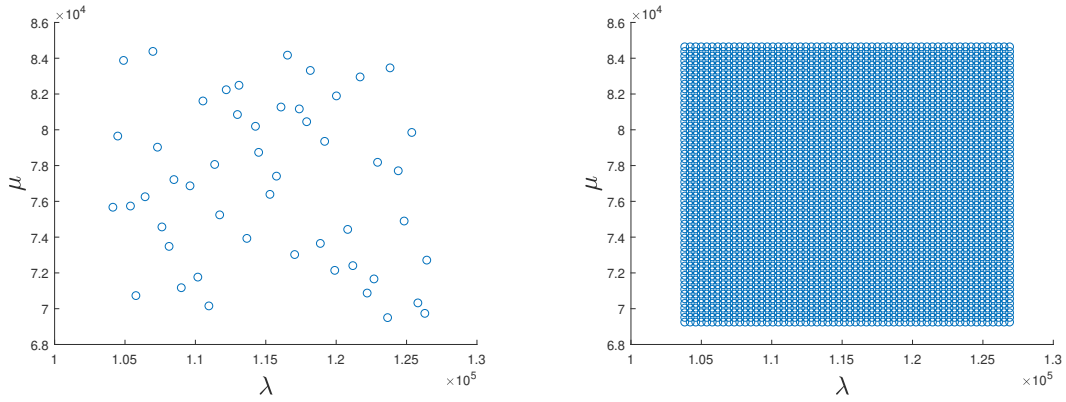


Figure 4.5 – 2D grid of points (a) Sample grid of points, N_s (b) Total number of grid of points, N_g

the RBF interpolation method.

1. To build the surrogate model, initially, we randomly choose(using LHS sampling) N_s sample grid points out of N_g number of grid points. We solve the FE model to calculate the displacement vectors at these N_s grid points (see Figure 4.5(a)). We obtain the N_s displacement vectors which we can store in the matrix $\boldsymbol{X}$ as

$$\boldsymbol{X} = [y_1, ..., y_{N_s}] \tag{4.2}$$

   The size of the matrix $\boldsymbol{X}$ is $N \times N_s$, where N is the size of the displacement column vector, y.

2. In the second step, we perform the singular value decomposition (SVD) of the matrix $\boldsymbol{X}$ as

$$SVD(\boldsymbol{X}) = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\boldsymbol{T}} \tag{4.3}$$

   where $\boldsymbol{U}$ is an $N \times N$ unitary matrix and its columns are called the left-singular vectors, $\boldsymbol{\Sigma}$ is an $N \times N_s$ diagonal matrix whose diagonal entries are known as the singular values of $\boldsymbol{X}$, $\boldsymbol{V}$ is an $N_s \times N_s$ unitary matrix and its columns are called the right-singular vectors. We use the left-singular vectors as basis vectors corresponding to the largest singular values to approximate the displacement vector, y, at any grid point x as

$$y(x) = \sum_{k=1}^{N_b} \alpha_k(x)\phi_k = [\boldsymbol{\phi}]\alpha \tag{4.4}$$

   where $\boldsymbol{\phi}$ is the matrix of the size, $N \times N_b$, N_b is the number of basis vectors such that $N_b \leq N_s$, $\alpha \in \mathbb{R}^{N_b}$ are the unknown coefficients.

3. In the next step, the idea is to find the values of the unknown α coefficients at chosen N_s sample grid points. As we know the values of the displacement vector, y, at chosen N_s sample grid points, thanks to the FE calculations, hence using Least-squares minimization problem, we can compute an unknown vector α of coefficients at any i^{th} sample grid point for $i = 1$ to N_s as

$$\alpha(x) = \underset{\alpha \in \mathbb{R}^{N_b}}{\text{Arg Min}}\, J(\alpha) \tag{4.5}$$

where $J(\alpha) = \|\sum_{k=1}^{N_b} \phi_k \alpha_k(x) - y(x)\|_2^2$. Then it is possible to compute displacement vector y at any sampled grid point x_i using Equation (4.4).

4. In order to calculate displacement at any grid point x as we have a total N_g number of grid points (see Figure 4.5(b)), we need to know the values of the coefficients α at any grid point x. For that, we use the RBF interpolation method. In this method, we compute the unknown vector at any grid point x from known vector α at N_s sample grid points. We approximate the row vector α of the size $1 \times N_b$ at any grid point x by defining radial basis function at N_s number of sampled grid points as

$$\alpha(x) = \sum_{k=1}^{N_b} \sum_{i=1}^{N_s} \varphi(\|x - x_i\|) w_{ik} = [\boldsymbol{\varphi}][w] \tag{4.6}$$

where φ is a radial basis function, $\boldsymbol{\varphi}$ is the matrix of size, $1 \times N_s$, and w is the matrix of the unknown weights with size $N_s \times N_b$. We choose Gaussian RBF function such that $\varphi(\|x - x_i\|) = exp(-(\frac{x-x_i}{2\boldsymbol{\sigma_{rbf}}})^2)$ and this function is parametrised by the shape parameter value, i.e, $\frac{1}{2\boldsymbol{\sigma_{rbf}}}$. The unknown weights w at any grid point x_i can be computed by solving following linear system of the equations:

$$[\boldsymbol{\varphi}_{global}][w] = [\alpha_{global}] \tag{4.7}$$

where $\boldsymbol{\varphi}_{global}$ is the matrix of size, $N_s \times N_s$ and which is given as

$$\begin{bmatrix} \varphi(\|x_1 - x_1\|) & \varphi(\|x_2 - x_1\|) & \dots & \varphi(\|x_{N_s} - x_1\|) \\ \varphi(\|x_1 - x_2\|) & \varphi(\|x_2 - x_2\|) & \dots & \varphi(\|x_{N_s} - x_2\|) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\|x_1 - x_{N_s}\|) & \varphi(\|x_2 - x_{N_s}\|) & \dots & \varphi(\|x_{N_s} - x_{N_s}\|) \end{bmatrix},$$

w is the matrix of size, $N_s \times N_b$ and which is given as

$$\begin{bmatrix} w_{11} & \dots & w_{1N_b} \\ w_{21} & \dots & w_{2N_b} \\ \vdots & \vdots & \vdots \\ w_{N_s1} & \dots & w_{N_sN_b} \end{bmatrix},$$

α_{global} is the matrix of size, $N_s \times N_b$ and which is given as

$$\begin{bmatrix} \alpha_1(x_1) & \dots & \alpha_{N_b}(x_1) \\ \alpha_1(x_2) & \dots & \alpha_{N_b}(x_2) \\ \vdots & \vdots & \vdots \\ \alpha_1(x_{N_s}) & \dots & \alpha_{N_b}(x_{N_s}) \end{bmatrix}.$$

We can solve Equation (4.7) to compute w by following equation

$$w = [\boldsymbol{\varphi}_{global}]^{-1}\alpha_{global} \tag{4.8}$$

We can compute the unknown row vector α at any grid point x by knowing weights w using Equation (4.6). If we know the values of α then we can compute the displacement vector, y, at any grid point x that helps to avoid solving FE calculations at each grid point.

Now, we apply the set-valued inverse method along with $GCONS$ outlier detection method (Algorithm 2.3.3.3) to identify the set of elastic parameters when there is random noise in the data. To generate the displacement measurement data $\tilde{y}$ (12096 measurements), exact displacement data y^{Ref} is simulated by a Finite Element (FE) model (6048 nodes, 12094 elements) as shown in Figure 3.2 considering the reference values $\lambda_0 = 1.15 \cdot 10^5 MPa$ and $\mu_0 = 7.69 \cdot 10^4 MPa$. We also consider a possible Gaussian white noise with 0 mean (no systematic bias) and with standard deviation σ. In the current work, σ was taken as 5% of the average of all the exact displacement values and in practical cases it can be assumed that σ can be deduced from the measurement technique. The information on the measurement $\tilde{y}$ was described in an interval form: $[\ \tilde{y} - 2\sigma,\quad \tilde{y} + 2\sigma\ ]$. Prior information about the parameters ($\mathbb{S}_\theta^0$) is considered as a uniform $2D$ box (see Figure 4.5(b)) $\lambda^p \times \mu^p$ with $\lambda^p = [1.03 \cdot 10^5,\ 1.26 \cdot 10^5]MPa$ and $\mu^p = [\ 6.92 \cdot 10^4,\ 8.46 \cdot 10^4]MPa$. For the surrogate modeling with RBF interpolation method, we have used $N_s = 50$, $N_g = 3600$, $\boldsymbol{\sigma_{rbf}} = 0.0293$. The value of $\boldsymbol{\sigma_{rbf}}$ plays important role in the accuracy of RBF surrogate modeling. We have optimized $\boldsymbol{\sigma_{rbf}}$ using L^2 norm between the displacement vector y computed using the FE model and the surrogate model,i.e., $\|y_{FE} - y_{surrogate}\|_2^2$ at one sample grid point. We selected $\boldsymbol{\sigma_{rbf}}$ which minimizes the

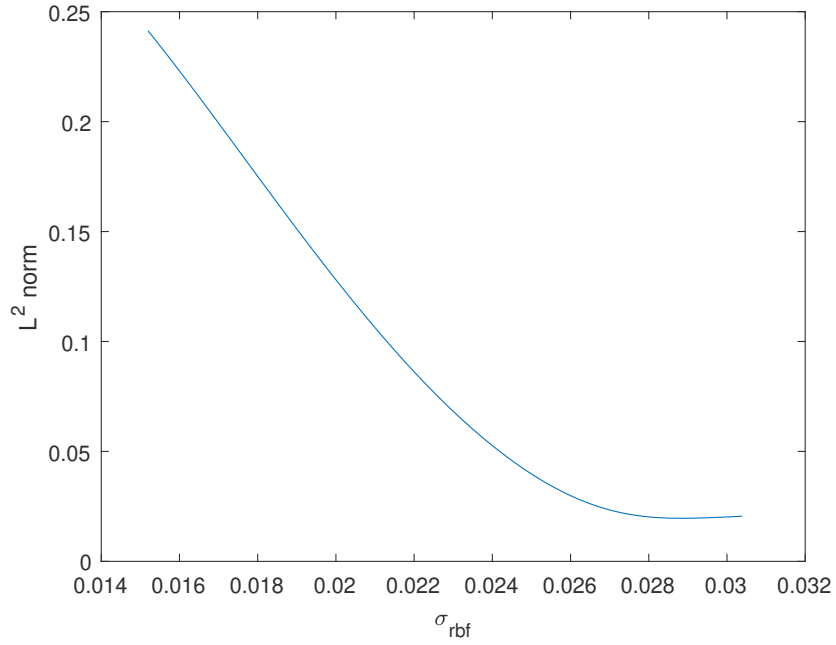


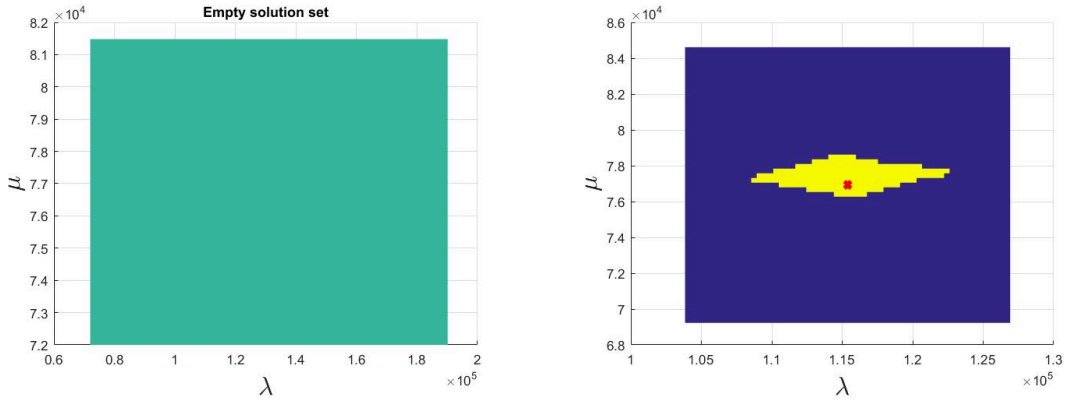
Figure 4.6 – L^2 norm $= \|y_{FEM} - y_{surrogate}\|_2^2$

L^2 norm as shown in the Figure 4.6.

Figure 4.7(a) shows that the identified set (green color) when taking all the measurements is empty due to inconsistency within the measurements. To obtain a non-empty solution set, we use our proposed solution and Algorithm 2.3.3.3 with the value of the $GCONS_{threshold}(\tau)$ settled to 0.1. We use a low value of GCONS to ensure that a high enough number of measurements will be included. Figure 4.7(b) shows the feasible set (yellow color) of the identified parameter using $GCONS$ method, with 1233 measurements removed. We can note here that the exact value of the parameter is included in the solution set (shown by red mark).

We solve this identification problem using a surrogate model based on the RBF interpolation method. It took a total time of 55 minutes to solve the problem on a standard laptop. If solved without surrogate modeling, it would have taken almost one day. Yet in the set-valued inverse method, computation time is divided into two parts; the first one corresponding to the solving of the FE model to compute displacement at each grid point, and the second part is related to the post-processing of a large number of measurements with $GCONS$ outlier detection algorithm. For the first part, we can use surrogate modeling to reduce computational time, and in this application, out of 55 minutes, it has taken 22 minutes. Hence, the $GCONS$ algorithm takes 33 minutes to process 12096 measurements to detect outliers.

One way to reduce the computation time with the $GCONS$ algorithm is that, generally, in the $GCONS$ algorithm, we start with the first two measurements



(a) (b)

Figure 4.7 – 2D identification with $GCONS$ method with a large number of measurements. (a) Empty solution set. (b) Solution set after detecting outlier

with the largest $GDOC$ values (line 2 of the Algorithm 2.3.3.3). To reduce the time when we have a large number of measurements, one way is to start with a subgroub of measurements whose $GDOC$ values are high so that it reduces the time for computation with $GCONS$ algorithm. In this application with the $GCONS$ algorithm, out of 12096 measurements, the initial subgroub is fixed to the first 8000 measurements with high $GDOC$ value. Then time taken by the $GCONS$ algorithm reduces to 14 minutes from 33 minutes (when starting with the first two measurements with high $GDOC$ value). The total time taken to solve the identification problem is 22+14=36 minutes. With this modification in the $GCONS$ algorithm, we did not notice the change in the final result of the solution set.

We also solved the sample problem of identification with a large FE model without surrogate modeling, and Figure 4.8 shows the result of the identified solution set. We can notice that both results(see Figure 4.7(b) with surrogate modeling and Figure 4.8 without surrogate modeling) include the true parameter value, but the solution set's shape is different. The possible reason may be due to the use of the surrogate model to solve the large FE model, but at the same time, we can reduce the computation time with surrogate modeling.

We provided the above application to show how we can solve the identification problem with a set-valued inverse problem when we have a large FE model using surrogate modeling. In our work, we also studied the same application even with a larger number of measurements, i.e., 37224 full-field displacement measurements. To do this, with the same problem, we increased the number of nodes to 18612 and number of elements to 37222 for the FE mesh. We solved the same identification problem with set-valued inverse method using the surrogate modeling. With the

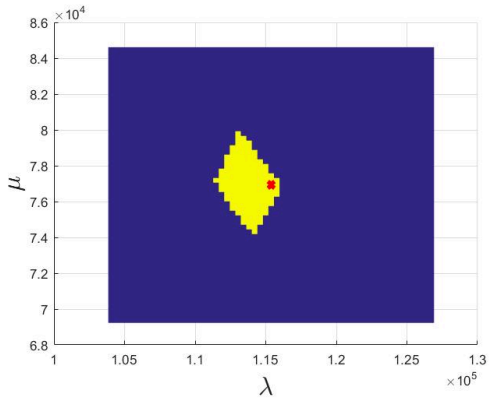


Figure 4.8 – Solution set without surrogate modeling

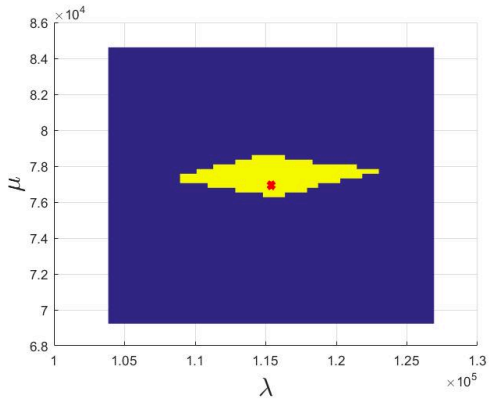
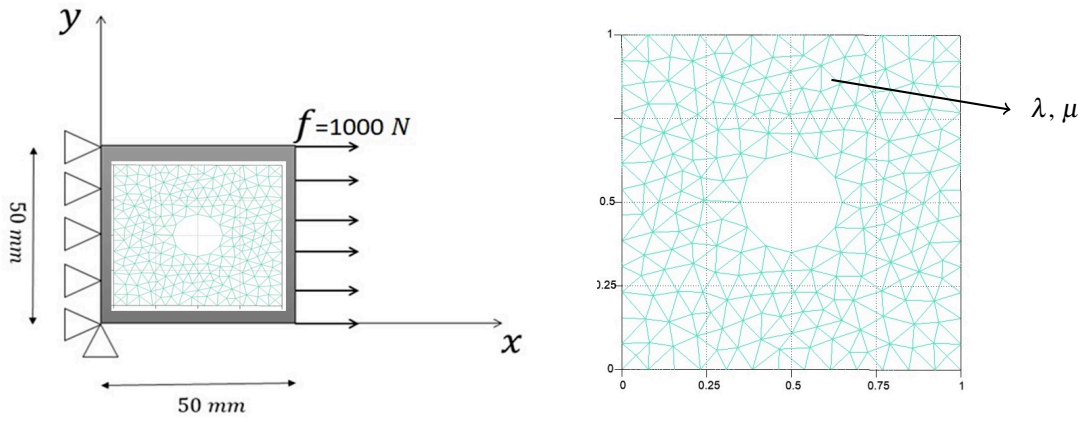


Figure 4.9 – 2D identification with $GCONS$ method with a large number of measurements (N=37224)

$GCONS$ outlier detection method and by setting the $GCONS_{threshold}(\tau)$ settled to 0.1, we obtained the solution set of the parameters as shown in Figure 4.9. To obtain non-empty solution set we removed 3515 inconsistent measurements. The above results showed how we could solve set-valued inverse problems using surrogate modeling, especially when we have a large FE model and large number of inconsistent measurements.

4.3 Application of the identification strategy to damage detection

Damage detection plays an important role in structural health monitoring (SHM) of mechanical or civil engineering structures [Del Grosso, 2013, Cawley, 2018]. The term SHM refers to the process of applying a damage detection and characterization strategy for engineering structures. Within the framework of damage detection in the material structures, there are various ways to define the criterion of damage



a) A plate with hole b) FEM mesh

Figure 4.10 – Homogeneous isotropic elastic structure

in the material [Sinou, 2009]. In our work, we use "a reduction in stiffness of a structure," as the definition of damage to indicate the presence of damage in the structure. There are two types of data, which are dynamical and static response data, which are useful for solving the inverse problem to detect damage to the structure [Truman and Terlaje, 2006]. When it comes to dynamical response data, it includes vibrational data [Doebling et al., 1998] that helps to identify modal parameters (natural frequencies, mode shapes, and modal damping). These modal parameters are functions of the structure's physical properties (mass, damping and stiffness (including elastic properties)). A criterion of damage in the material using dynamical response data is that the reduction in natural frequencies leads to a decrease in the structure's stiffness, which helps detect damage in the material. The static response data includes displacement or strain data [Truman and Terlaje, 2006]. The displacements are the function of the structure's overall stiffness. Due to the static nature of the experimental test [Truman and Terlaje, 2006, Augusto S. Terlaje and Truman, 2007], identification methods use static response data with inverse methods to detect damage in the material to check whether the structure stiffness is reduced or not.

In our work, we use static response data along with a set-valued inverse method to determine the location of damage in a homogeneous isotropic elastic structure. In this application, the measure we choose is that if there is a reduction in values of elastic parameters at a given location from its initial values, then that location is considered as a damaged part of the structure.

A square plate with a hole is chosen to represent such a homogeneous isotropic elastic structure in Figure 4.10. The plate is clamped on the left side and loaded on the right side by a uniform traction $f = 1000\ N/m$. To create the measurements that

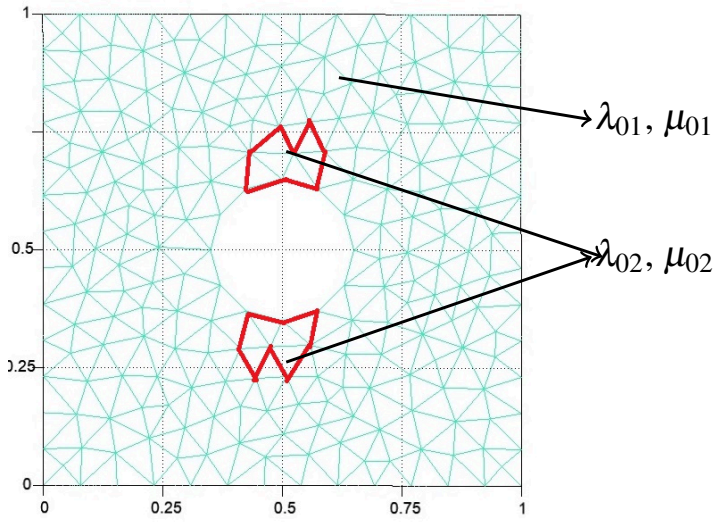
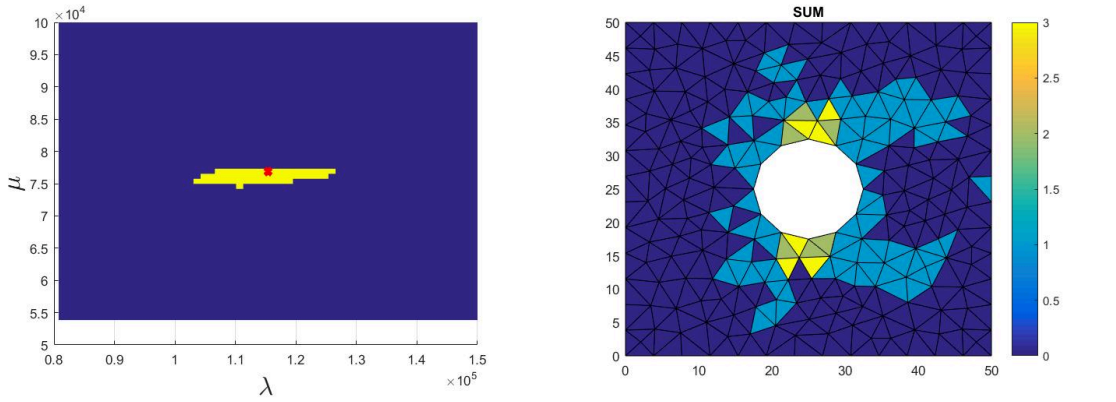


Figure 4.11 – A model to create measurements which are affected due to damage

are affected by the damage in the material, we first generate through simulation exact displacement data y^{Ref} by a Finite Element (FE) model (236 nodes, 408 elements) as shown in Figure 4.11. To do this, we consider the reference/exact values $\lambda_{01} = 1.15 \cdot 10^5 MPa$ and $\mu_{01} = 7.69 \cdot 10^4 MPa$ for the region outside the red boundaries and $\lambda_{02} = 9.23 \cdot 10^4 MPa$ and $\mu_{02} = 6.15 \cdot 10^4$ MPa for the area inside the red boundaries such that $\lambda_{02} < \lambda_{01}$ and $\mu_{02} < \mu_{01}$. To detect the damage, we use strain measurements instead of displacement. We derive strain from displacements such that each element of the FE mesh consists of three strain measurements, ε_{xx}, ε_{xy}, and ε_{yy} respectively. Hence, we have a total of 1224 strain measurements, as there are 408 elements. We also consider a possible Gaussian white noise with 0 mean (no systematic bias) and with standard deviation σ. In the current work, σ was taken as 5% of the average of all the exact strain values. We represent each strain measurements in xx direction in their interval form such that $[\tilde{\varepsilon}_{xx} - 3\sigma_1,\ \tilde{\varepsilon}_{xx} + 3\sigma_1]$ where $\sigma_1 = 5$ % of the average of all the exact strain values in the xx direction. Similarly, for the xy and yy direction, we define each stain measurement in its interval form as: $[\tilde{\varepsilon}_{xy} - 3\sigma_2,\ \tilde{\varepsilon}_{xy} + 3\sigma_2]$, $[\tilde{\varepsilon}_{yy} - 3\sigma_3,\ \tilde{\varepsilon}_{yy} + 3\sigma_3]$. Prior information about the parameters ($\mathbb{S}_\theta^0$) is considered as a $2D$ box $\lambda^p \times \mu^p$ with $\lambda^p = [0.8\ 10^5,$ $1.5\ 10^5] MPa$ and $\mu^p = [\ 5.38\ 10^4,\ 1\ 10^5] MPa$.



a) Identification of parameters b) Identification of the damage location

Figure 4.12 – Damage detection

4.3.1 Damage detection when there is no noise in the strain measurement data

We first apply the set-valued inverse method to identify the set of elastic parameters when there is no noise in the data. The measurement data was chosen such that $\tilde{\varepsilon} = \varepsilon^{Ref}$, and the information on the measurement $\tilde{\varepsilon}$ was described in an interval form: $[\tilde{\varepsilon}_{xx} - 3\sigma_1,\ \tilde{\varepsilon}_{xx} + 3\sigma_1]$, $[\tilde{\varepsilon}_{xy} - 3\sigma_2,\ \tilde{\varepsilon}_{xy} + 3\sigma_2]$, $[\tilde{\varepsilon}_{yy} - 3\sigma_3,\ \tilde{\varepsilon}_{yy} + 3\sigma_3]$.

To detect damage location, we apply the proposed strategy to identify the elastic parameters of an isotropic structure from full-field strain measurements affected due to damage and detect outliers in the measurements assuming the material model is still isotropic homogeneous. Figure 4.12(a) shows the identified solution set of parameters with 116 strain measurements removed out of 1224. As in this particular case, we do not have measurement error, we can conclude that the 116 removed strain measurements are due to model error, i.e., due to damage. Thanks to our $GCONS$ outlier detection method, we can detect outliers, which are due to measurement or model error. We determine the location of removed strain measurement, i.e., to which mesh element they belong. A maximum of 3 outliers can be present on each element as each element consists of 3 strain measurements. For that, we use a colormap (see Figure 4.12(b)) that maps the sum (number of outliers present on the element) from 0 to 3 over each mesh element. To determine the location of measurements that are outlier due to model error, the criterion we choose is that the sum over element should be greater than and equal to 2, i.e., each element contains two or three removed measurements present on them. The damaged area corresponds to the elements that contain two or three outliers present on them and these elements should be spatially connected with one another. We found such 12 elements that determine the location of the damaged part of the

structure as shown in the Figure 4.12(b). Indeed, these are 12 elements that are inside the red boundaries, as shown in Figure 4.11 and where we changed material parameter values such that $\lambda_{02} < \lambda_{01}$ and $\mu_{02} < \mu_{01}$. Here, we need to note that the identified solution set in Figure 4.12(a) corresponds to the undamaged part of the material as we removed outliers due to model error. The solution set still contains the reference values (λ_{01}, μ_{01}) denoted by the red mark corresponding to the undamaged or homogeneous part of the material.

4.3.2 Damage detection when there is a random noise in the strain measurement data

We apply the set-valued inverse method to identify the set of elastic parameters when there is a random noise in the data. The measurement data $\tilde{\varepsilon}$ is created from ε^{Ref} by adding to it a Gaussian white noise with standard deviation σ, and the information on the measurement $\tilde{\varepsilon}$ was described in an interval form: $[\tilde{\varepsilon}_{xx} - 3\sigma_1,\ \tilde{\varepsilon}_{xx} + 3\sigma_1]$, $[\tilde{\varepsilon}_{xy} - 3\sigma_2,\ \tilde{\varepsilon}_{xy} + 3\sigma_2]$, $[\tilde{\varepsilon}_{yy} - 3\sigma_3,\ \tilde{\varepsilon}_{yy} + 3\sigma_3]$.

Figure 4.13(a) shows the identified solution set of parameters with 275 strain measurements removed out of 1224 using $GCONS$ outlier detection algorithm. The solution set still contains the reference values (λ_{01}, μ_{01}) denoted by the red mark corresponding to the undamaged or homogeneous part of the material. As in this case, we do have both measurement error and model error, the 275 removed strain measurements may be due either to measurement or model error, i.e., due to damage. We determine the location of removed strain measurement, i.e., to which mesh element they belong. A maximum of 3 outliers can be present on each element as each element consists of 3 strain measurements. For that, we use a colormap (see Figure 4.13(b)) that maps the sum (number of outliers present on the element) from 0 to 3 over each mesh element. We use a random noise to create the measurement error; hence the location of the removed measurement that are affected due to measurement error should not display spatial coherence. Such a fact is illustrated in Figure 4.13(b) where we can see a few elements which contain two removed measurement, but are not connected to each other. Then how can one detect the damaged location when we have inconsistencies due to both measurement error and model error? The answer is to find elements that contain two or three removed measurement, and are spatially connected with one another. By using this criterion, we choose 17 such elements (elements with yellow color) which represent the approximated damaged zone in the material, as shown in Figure 4.13(c). These 17 elements contain 51 strain measurements, and the intersection of all the solution sets corresponding to those measurements gives a

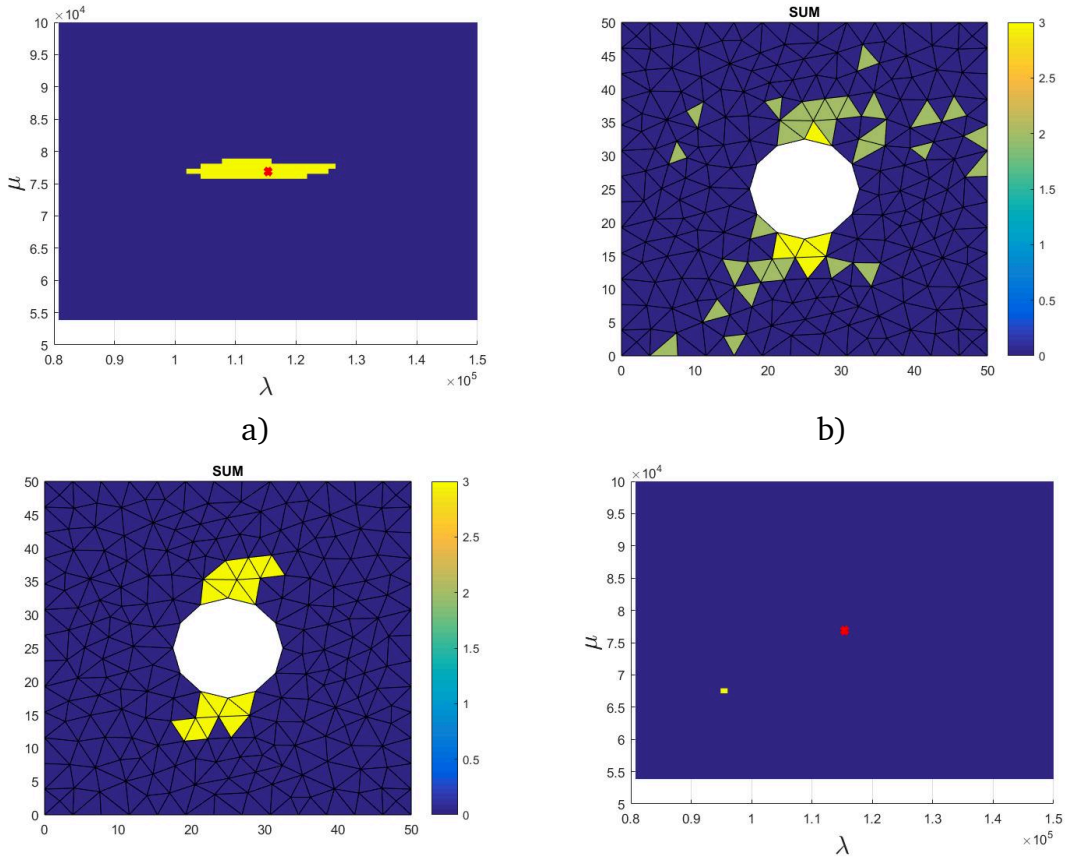


Figure 4.13 – Damage detection when there is a noise in the data: a) Identification of parameters, b) Location of removed measurements c) approximated damage location, d) identification from the approximated damaged part.

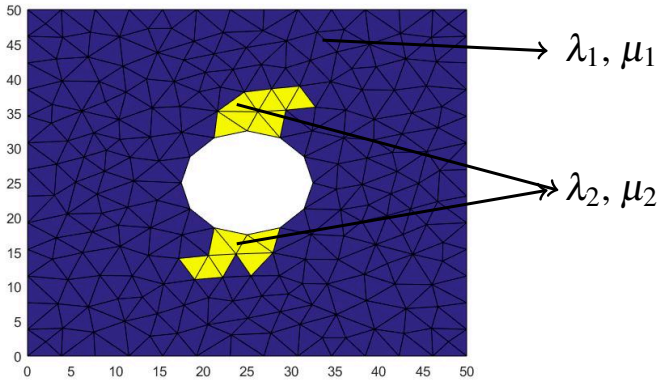
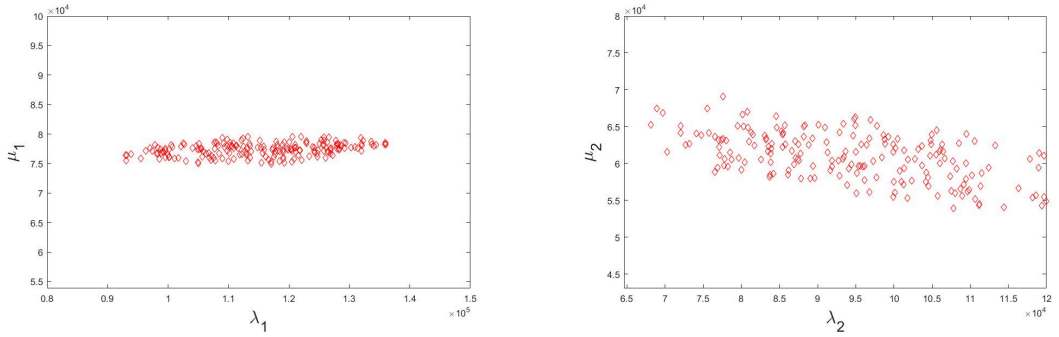


Figure 4.14 – Heterogeneous material with approximated damaged zone (after damage detection)

solution set of parameters (yellow color), shown in Figure 4.13(d). Though small, this solution set does not contain the reference value (λ_{01}, μ_{01}) denoted by the red mark, which corresponds to the material's undamaged or homogeneous part. We can see that the value of the parameter (yellow mark in Figure 4.13(d) is decreased compared to the parameter values that correspond to the identification from the undamaged part (see Figure 4.13 (a)), and this also confirms the damage in the material, i.e, the values of the parameters λ, μ (damaged part) in Figure 4.13(d) are close to reference values (λ_{02}, μ_{02}) such that $\lambda_{02} < \lambda_{01}$ and $\mu_{02} < \mu_{01}$. This method to validate the approximate damage zone may not work well because of the hypothesis that the material remains homogeneous even when considering 51 strain measurements.

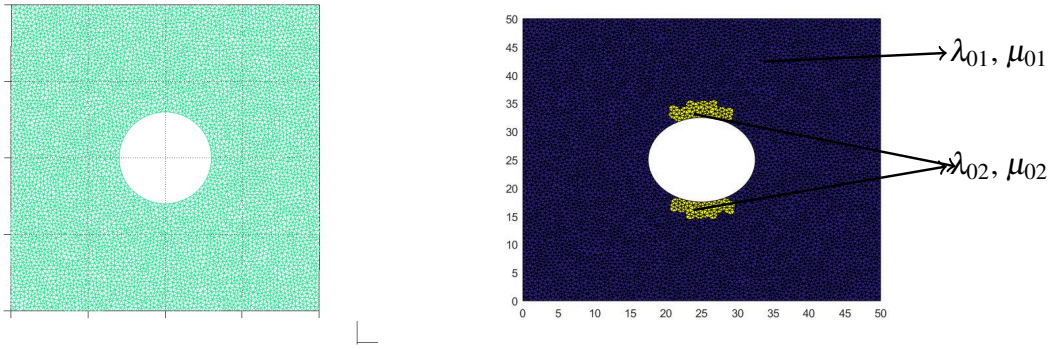
Hence, for this particular application, another way to validate the approximated damaged zone is to perform a 4D parameter identification, i.e., λ_1, μ_1 represents the parameters for the undamaged part, and λ_2, μ_2 represents the parameters for the damaged part from strain measurements affected due to damage (see Figure 4.14). The idea is to check whether λ_2 is less than λ_1, and μ_2 is less that μ_1 because a reduction in the values of the material properties confirms the damage in the material.

We performed 4D identification with the same strain measurement data, which is used to detect damage in the material. Here, the only difference is that (see Figure 4.14) material has become heterogeneous, i.e., $(\lambda_1 \neq \lambda_2)$ and $(\mu_1 \neq \mu_2)$, thanks to the approximated location of the damage. Prior information about the parameters $(\mathbb{S}_\theta^0)$ is considered as a non-uniform $4D$ box of grid of points $\lambda_1^p \times \mu_1^p \times \lambda_2^p \times \mu_2^p$ with $\lambda_1^p = [0.8 \cdot 10^5, 1.5 \cdot 10^5]MPa$, $\lambda_2^p = [0.64 \cdot 10^4, 1.2 \cdot 10^4]MPa$



(a) (b)

Figure 4.15 – 4D identification with $GCONS$ method. (a) Solution set($\lambda_1 \times \mu_1$). (b) Solution set ($\lambda_2 \times \mu_2$).



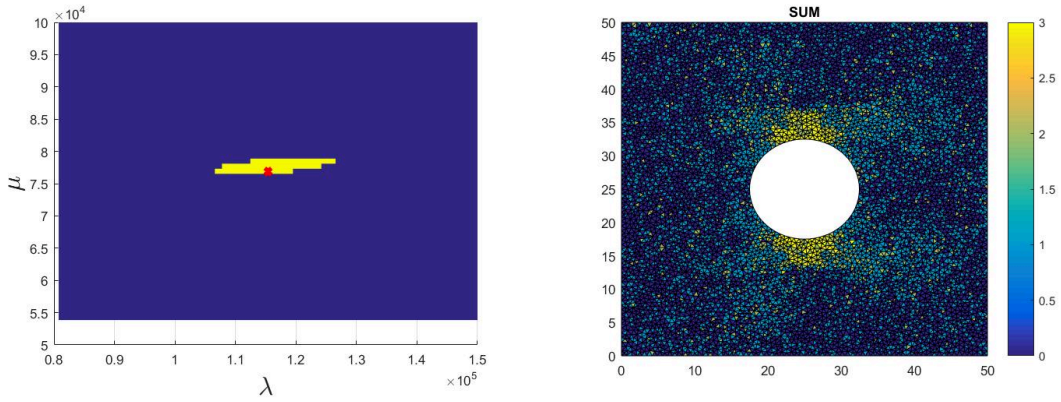
a) Plate with hole b) FE mesh

Figure 4.16 – A model to create large number of measurements which are affected due to damage

and $\mu_1^p = [\,5.38 \cdot 10^4,\, 10 \cdot 10^4]MPa$, $\mu_2^p = [\,4.3 \cdot 10^4,\, 8 \cdot 10^4]MPa$. With LHS sampling method, to describe $4D$ box of non-uniform grid of points, we used 20000 points for the grid.

We solved the inverse problem with the $GCONS$ outlier detection method, as usual, with the value of the $GCONS_{threshold}(\tau)$ settled to 0.1. The corresponding identified set of 4 parameter is presented in Figure 4.15 , where 319 measurements were removed. The solution set still contains the reference values $\lambda_{01} = 1.15 \cdot 10^5 MPa$, $\mu_{01} = 7.69 \cdot 10^4 MPa$ and $\lambda_{02} = 9.23 \cdot 10^4 MPa$, $\mu_{02} = 6.15 \cdot 10^4 MPa$ and these are the values we used to create data affected due to damage. This 4D identification confirms that $\lambda_2 < \lambda_1$ and $\mu_2 < \mu_1$.

In our work, we also studied the same application of damage detection, even with a large number of measurements, i.e., 37824 full-field strain measurements. With the same problem, we increased the number of nodes to 6482 and the number of elements to 12608 for the FE mesh 4.16(a). We solved the identification problem



a) Solution set after outlier detection b) Damage identification

Figure 4.17 – Damage detection with large number of measurements

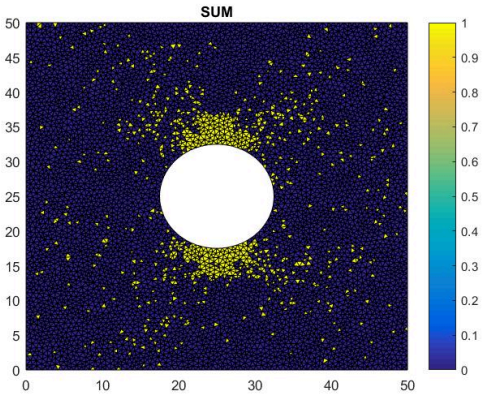


Figure 4.18 – Damage identification (location of outliers due to measurement and model error)

with the set-valued inverse method using surrogate modeling. We created measurement data affected due to damage in the same way as before (see Figure 4.16(b)) such that $\lambda_{02} < \lambda_{01}$ and $\mu_{02} < \mu_{01}$.

With the $GCONS$ outlier detection method and by setting the $GCONS_{threshold}(\tau)$ settled to 0.1, we obtained the solution set of the parameters as shown in Figure 4.17(a). To obtain a non-empty solution set, we removed 6522 inconsistent measurements. Thanks to the $GCONS$ outlier detection method, we could differentiate between outliers due to model and measurement error. Like in the previous example, we can also see in Figure 4.17(b) a few elements contain 2 or 3 outliers present on them. We can approximate damage location by finding elements that contain two or three outliers present on them, and that are connected with one another as shown in Figure 4.18. It should be noted that Figures 4.17(b) and 4.18 are the same; in order to visualize clearly, we transformed the sum 2 and 3 in Figure 4.17(b) and sum 1 in Figure 4.17(b) in to 1 and 0 respectively. In this

application, we have shown that how the proposed strategy can be useful when we need to detect damage in the material.

4.4 Summary

In this chapter, we presented the set-valued inverse method applications when identifying high dimensional data and damage detection in the material structure. In the case of identifying a large number of parameters or with a large number of measurements, we can face a too important computational complexity with the set-valued inverse method. To solve such complexity, firstly, we presented a non-uniform description of a grid of points using Latin hypercube sampling when we have to identify a large number of parameters. We also demonstrated the use of surrogate modeling along with the set-valued inverse method when FE model is large or measurements are in a large number. The results also showed how we could deal with a large number of inconsistent measurements using the outlier detection method. The second application we considered is damage detection in the material structure. With this application, we showed how we could differentiate between outliers due to measurement and model error. To conclude, we say that interestingly, the set-valued inverse method does not only help to obtain a feasible set of parameters and detect outliers but extend its capability to identify the nature of outliers in the measurements.

Conclusion

In this thesis, an efficient inverse problem strategy has been developed based on a non-probabilistic approach to uncertainty representation. Our strategy uses sets or intervals within the non-probabilistic uncertainty modeling approach to represent uncertainty on measurements and prior information on parameters. Our methodology has proved effective for estimating material property parameters from full-field displacement measurements, especially when dealing with inconsistent measurements or outliers. Unlike other set-based identification strategies such as Q-intersection method [Drevelle and Bonnifait, 2012], the proposed strategy is efficient when dealing with inconsistent measurements in the inverse problem when they are in large amounts. The identification strategy makes it possible to solve critical application in structural damage detection, to estimate the location of the damage in the material.

The identification strategy makes use of the information such as measurement and prior information about parameters. The interesting thing about this strategy is that the amount of needed assumptions about uncertainty is minimal, as we use intervals. The prior information about parameters is described through hypercubes, i.e., a multidimensional extension of intervals. In this set-valued inverse strategy, we use a grid of points to describe the hypercube or sets numerically. Thanks to set theory, we introduced indicators of consistency of measurements, using them to propose selection methods to detect inconsistent or outlier measurements. Unlike other interval-valued inverse methods, this strategy does not require to propagate intervals through the model while solving the inverse problem.

The following numerical applications have been studied:

1. identification of the elastic properties of homogeneous isotropic material with the proposed outlier detection methods;

2. the comparison of the set-valued inverse method with the Bayesian inference method in two situations: one where the Bayesian model was accurate, the other where it was misspecified;

3. identification with higher dimensional space;

4. identification of damage location in the material structure.

Our main conclusions from applications 1 and 2 are:

- The Bayesian and set-valued approaches deliver solutions of similar quality when the former makes the right assumptions. Still, the Bayesian approach is quite sensitive to misspecification and the presence of outliers, while our approaches are much more robust.

- When selecting measurements, it seems essential to integrate possible interactions between the removed measurements to ensure a good identification, especially when the presence of outliers of anomalies is suspected.

With application 3 that concerns higher dimensional space, we have shown how using a non-uniform description of a grid of points to describe prior information about parameters and surrogate modeling helps to solve computational complexity with the set-valued inverse method. Application 4 demonstrates how we could differentiate between inconsistent measurements due to measurement and model error to detect the material's damage location.

To conclude, the set-valued inverse method helps obtain a feasible set of parameters and detect inconsistent or outlier measurements , and can also be instrumental to identify the nature of outliers in the measurements. The presented strategy is generic to any inverse problem, especially when we have incomplete information to solve the inverse problem .

In future work, the following aspects would be interesting to investigate:

- The application of strategy to non-linear material models.

- Composite materials are heterogeneous in nature, and their characterization requires analysis at different scales, i.e., micro and macro. In the context of damage detection for composite structure, it would be interesting to know that one could identify subgroups of consistent measurements at different scales that are not consistent between them, hence potentially identifying damaged zones that behave differently.

- Experimental validation of proposed outlier detection methods.

- Is it possible to integrate introduced indicators of consistency measures with other interval-valued approaches or non-probabilistic uncertainty methods while solving an inverse problem?

- Application of outlier detection methods to fault detection problems.

References

[Akpan et al., 2001] Akpan, U., Koko, T., Orisamolu, I., and Gallant, B. (2001). Practical fuzzy finite element analysis of structures. *Finite Elements in Analysis and Design*, 38:93–111, doi: http://dx.doi.org/10.1016/S0168-874X(01)00052-X.

[Anderson, 1976] Anderson, G. (1976). Error propagation by the monte carlo method in geochemical calculations. *Geochimica et Cosmochimica Acta*, 40(12):1533 – 1538, doi: http://dx.doi.org/https://doi.org/10.1016/0016-7037(76)90092-2.

[Augusto S. Terlaje and Truman, 2007] Augusto S. Terlaje, I. and Truman, K. Z. (2007). Parameter identification and damage detection using structural optimization and static response data. *Advances in Structural Engineering*, 10(6):607–621, doi: http://dx.doi.org/10.1260/136943307783571409.

[Bajer and Holena, 2012] Bajer, L. and Holena, M. (2012). Rbf-based surrogate model for evolutionary optimization. In *ITAT*.

[Batou and Soize, 2013] Batou, A. and Soize, C. (2013). Stochastic modeling and identification of an uncertain computational dynamical model with random fields properties and model uncertainties. *Archive of Applied Mechanics*, 83, doi: http://dx.doi.org/10.1007/s00419-012-0720-7.

[Baudrit and Dubois, 2006] Baudrit, C. and Dubois, D. (2006). Practical representations of incomplete probabilistic knowledge. *Computational Statistics & Data Analysis*, 51(1):86 – 108, doi: http://dx.doi.org/https://doi.org/10.1016/j.csda.2006.02.009.

[Blais, 2010] Blais, J. A. R. (2010). Least squares for practitioners. *Mathematical Problems in Engineering*, 2010:19.

[Braems et al., 2001] Braems, I., Berthier, F., Jaulin, L., Kieffer, M., and Walter, E. (2001). Guaranteed estimation of electrochemical parameters by set inversion using interval analysis. *Journal of Electroanalytical Chemistry*, 495.

[Cawley, 2018] Cawley, P. (2018). Structural health monitoring: Closing the gap between research and industrial deployment. *Structural Health Monitoring: An International Journal*, page 147592171775004, doi: http://dx.doi.org/10.1177/1475921717750047.

[Chen et al., 2002] Chen, S., Lian, H., and Yang, X. (2002). Interval static displacement analysis for structures with interval parameters. *International Journal for Numerical Methods in Engineering*, 53(2):393–407, doi: http://dx.doi.org/10.1002/nme.281.

[Chen et al., 2000] Chen, Y., Breen, P. A., and Andrew, N. L. (2000). Impacts of outliers and mis-specification of priors on bayesian fisheries-stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(11):2293–2305, doi: http://dx.doi.org/10.1139/f00-208.

[Chetih and Messali, 2015] Chetih, N. and Messali, Z. (2015). Tomographic image reconstruction using filtered back projection (fbp) and algebraic reconstruction technique (art). In *2015 3rd International Conference on Control, Engineering Information Technology (CEIT)*, pages 1–6.

[Couto et al., 2013] Couto, P., Damasceno, J., and Pinheiro de Oliveira, S. (2013). *Monte Carlo Simulations Applied to Uncertainty in Measurement*, pages 27–51.

[Cunha et al., 2014] Cunha, A., Nasser, R., Sampaio, R., Lopes, H., and Breitman, K. (2014). Uncertainty quantification through the monte carlo method in a cloud computing setting. *Computer Physics Communications*, 185(5):1355 – 1363, doi: http://dx.doi.org/https://doi.org/10.1016/j.cpc.2014.01.006.

[Dashti and Stuart, 2017] Dashti, M. and Stuart, A. M. (2017). *The Bayesian Approach to Inverse Problems*, pages 311–428. Springer International Publishing, Cham.

[De Martino and De Martino, 2018] De Martino, A. and De Martino, D. (2018). An introduction to the maximum entropy approach and its application to inference problems in biology. *Heliyon*, 4(4):e00596, doi: http://dx.doi.org/https://doi.org/10.1016/j.heliyon.2018.e00596.

[Del Grosso, 2013] Del Grosso, A. (2013). Structural health monitoring: research and practice.

[Doebling et al., 1998] Doebling, S., Farrar, C., and Prime, M. (1998). A summary review of vibration-based damage identification methods. *The Shock and Vibration Digest*, 30:91–105, doi: http://dx.doi.org/10.1177/058310249803000201.

[Dong and Shah, 1987] Dong, W. and Shah, H. C. (1987). Vertex method for computing functions of fuzzy variables. *Fuzzy Sets and Systems*, 24(1):65 – 78, doi: http://dx.doi.org/https://doi.org/10.1016/0165-0114(87)90114-X.

[Drevelle and Bonnifait, 2012] Drevelle, V. and Bonnifait, P. (2012). Interval-based fault detection and identification applied to global positioning. *IFAC Proceedings Volumes*, 45(16):1085–1090.

[Eller and Valdivia, 2009] Eller, M. and Valdivia, N. P. (2009). Acoustic source identification using multiple frequency information. *Inverse Problems*, 25(11):115005.

[Emery et al., 2016] Emery, J., Grigoriu, M., and Field Jr., R. (2016). Bayesian methods for characterizing unknown parameters of material models. *Applied Mathematical Modelling*, 40(13):6395 – 6411, doi: http://dx.doi.org/https://doi.org/10.1016/j.apm.2016.01.046.

[Engl and Ramlau, 2015] Engl, H. W. and Ramlau, R. (2015). *Regularization of Inverse Problems*, pages 1233–1241. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Faes et al., 2019] Faes, M., Broggi, M., Patelli, E., Govers, Y., Mottershead, J., Beer, M., and Moens, D. (2019). A multivariate interval approach for inverse uncertainty quantification with limited experimental data. *Mechanical Systems and Signal Processing*, 118:534–548, doi: http://dx.doi.org/10.1016/j.ymssp.2018.08.050.

[Faes and Moens, 2019] Faes, M. and Moens, D. (2019). Recent trends in the modeling and quantification of non-probabilistic uncertainty. *Archives of Computational Methods in Engineering*, pages 1–39, doi: http://dx.doi.org/10.1007/s11831-019-09327-x.

[Fang et al., 2018] Fang, H., Gong, C., Li, C., Li, X., Su, H., and Gu, L. (2018). A surrogate model based nested optimization framework for inverse problem considering interval uncertainty. *Structural and Multidisciplinary Optimization*, 58(3):869–883.

[Fedele et al., 2012] Fedele, F., Muhanna, R. L., and Xiao, N. (2012). Interval-based inverse problems with uncertainties. *IFAC Proceedings Volumes*, 45(16):1079 – 1084, doi: http://dx.doi.org/https://doi.org/10.3182/20120711-3-BE-2027.00319. 16th IFAC Symposium on System Identification.

[Feeley et al., 2004] Feeley, R., Seiler, P., Packard, A., and Frenklach, M. (2004). Consistency of a reaction dataset. *The Journal of Physical Chemistry A*, 108(44):9573–9583.

[Ferson and Ginzburg, 1996] Ferson, S. and Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2):133 – 144, doi: http://dx.doi.org/https://doi.org/10.1016/S0951-8320(96)00071-3.

[Fish and Belytschko, 2007] Fish, J. and Belytschko, T. (2007). A first course in finite elements.

[Garloff, 2008] Garloff, J. (2008). Interval gaussian elimination with pivot tightening. *SIAM J. Matrix Analysis Applications*, 30:1761–1772, doi: http://dx.doi.org/10.1137/080729621.

[Gogu et al., 2010] Gogu, C., Haftka, R., Le Riche, R., Molimard, J., and Vautrin, A. (2010). Introduction to the bayesian approach applied to elastic constants identification. *AIAA Journal*, 48(5):893–903.

[Gong et al., 2020] Gong, W., Zhao, C., Juang, C. H., Tang, H., Wang, H., and Hu, X. (2020). Stratigraphic uncertainty modelling with random field approach. *Computers and Geotechnics*, 125:103681, doi: http://dx.doi.org/https://doi.org/10.1016/j.compgeo.2020.103681.

[Grédiac and Hild, 2012] Grédiac, M. and Hild, F. (2012). *Full-Field Measurements and Identification in Solid Mechanics*.

[Hadamard, 1902] Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princet University Bulletin*, 13:49–52.

[Halton, 1960] Halton, J. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.*, 2:84–90, doi: http://dx.doi.org/10.1007/BF01386213.

[Hansen, 1992] Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4):561–580, doi: http://dx.doi.org/10.1137/1034115.

[Harremoës and Topsoe, 2001] Harremoës, P. and Topsoe, F. (2001). Maximum entropy fundamentals. *Entropy*, 3:191–226, doi: http://dx.doi.org/10.3390/e3030191.

[Hegde et al., 2018] Hegde, A., Li, W., Oreluk, J., Packard, A., and Frenklach, M. (2018). Consistency analysis for massively inconsistent datasets in bound-to-bound data collaboration. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):429–456.

[Helton et al., 2004] Helton, J., Johnson, J., and Oberkampf, W. (2004). An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering System Safety*, 85(1):39 – 71, doi: http://dx.doi.org/https://doi.org/10.1016/j.ress.2004.03.025. Alternative Representations of Epistemic Uncertainty.

[Helton and Johnson, 2011] Helton, J. C. and Johnson, J. D. (2011). Quantification of margins and uncertainties: Alternative representations of epistemic uncertainty. *Reliability Engineering & System Safety*, 96(9):1034 – 1052, doi: http://dx.doi.org/https://doi.org/10.1016/j.ress.2011.02.013. Quantification of Margins and Uncertainties.

[Hora, 1996] Hora, S. C. (1996). Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management. *Reliability Engineering System Safety*, 54(2):217 – 223, doi: http://dx.doi.org/https://doi.org/10.1016/S0951-8320(96)00077-4. Treatment of Aleatory and Epistemic Uncertainty.

[Hose and Hanss, 2019] Hose, D. and Hanss, M. (2019). Consistent inverse probability and possibility propagation.

[Jaulin et al., 2001] Jaulin, L., Kieffer, M., Didrit, O., and Walter, E. (2001). *Applied Interval Analysis with Examples in Parameter and State Estimation, Robust Control and Robotics*.

[Jaulin and Walter, 1993] Jaulin, L. and Walter, E. (1993). Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica*, 29(4):1053 – 1064, doi: http://dx.doi.org/https://doi.org/10.1016/0005-1098(93)90106-4.

[Jaynes, 2003] Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

[Kaplan, 2021] Kaplan, D. (2021 (accessed January 1, 2021)). *Knee Point*. url: https://www.mathworks.com/matlabcentral/fileexchange/35094-knee-point.

[Kindermann and Raik, 2020] Kindermann, S. and Raik, K. (2020). A simplified l-curve method as error estimator. *ETNA - Electronic Transactions on Numerical Analysis*, 53:217–238, doi: http://dx.doi.org/10.1553/etna_vol53s217.

[Kong et al., 2019] Kong, M., Li, D., and Zhang, D. (2019). Research on the application of improved least square method in linear fitting. *IOP Conference Series: Earth and Environmental Science*, 252:052158.

[McKay et al., 1979] McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, doi: http://dx.doi.org/10.1080/00401706.1979.10489755.

[Mierlo et al., 2019] Mierlo, C., Faes, M., and Moens, D. (2019). Identification of visco-plastic material model parameters using interval fields.

[Moens and Hanss, 2011] Moens, D. and Hanss, M. (2011). Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: Recent advances. *Finite Elements in Analysis and Design*, 47(1):4 – 16.

[Moore, 1966] Moore, R. E. (1966). Interval analysis. *Prentice-Hall,Englewood Cliffs, New Jersey*, page 145.

[Muhanna and Mullen, 2001] Muhanna, R. L. and Mullen, R. L. (2001). Uncertainty in mechanics problems—interval–based approach. *Journal of Engineering Mechanics*, 127(6):557–566, doi: http://dx.doi.org/10.1061/(ASCE)0733-9399(2001)127:6(557).

[Müller, 2013] Müller, U. K. (2013). Risk of bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica*, 81(5):1805–1849.

[Neumaier, 1990] Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge.

[Ota et al., 2019] Ota, Y., Jiang, Y., Nakamura, G., and Uesaka, M. (2019). Bayesian inference approach to inverse problems in a financial mathematical model. *International Journal of Computer Mathematics*, 0(0):1–15, doi: http://dx.doi.org/10.1080/00207160.2019.1671978.

[Ozcanan and Atahan, 2020] Ozcanan, S. and Atahan, A. O. (2020). Radial basis function surrogate model-based optimization of guardrail post embedment depth in different soil conditions. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 234(2-3):739–761, doi: http://dx.doi.org/10.1177/0954407019848548.

[Peters and Ranson, 1982] Peters, W. H. and Ranson, W. F. (1982). Digital Imaging Techniques In Experimental Stress Analysis. *Optical Engineering*, 21(3):427 – 431, doi: http://dx.doi.org/10.1117/12.7972925.

[Saliby and Pacheco, 2003] Saliby, E. and Pacheco, F. (2003). An empirical evaluation of sampling methods in risk analysis simulation: Quasi-monte carlo, descriptive sampling, and latin hypercube sampling. volume 2, pages 1606– 1610 vol.2.

[Sandretto et al., 2014] Sandretto, D., Alexandre, J., Trombettoni, G., Daney, D., and Chabert, G. (2014). Certified calibration of a cable-driven robot using interval contractor programming. In *Computational Kinematics*, pages 209–217. Springer.

[Shafer, 1976] Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, N.J: Princeton University Press.

[Shary, 2001] Shary, S. (2001). Interval gauss-seidel method for generalized solution sets to interval linear systems. *Reliable Computing*, 7:141–155, doi: http://dx.doi.org/10.1023/A:1011422215157.

[Shields and Zhang, 2016] Shields, M. D. and Zhang, J. (2016). The generalization of latin hypercube sampling. *Reliability Engineering System Safety*, 148:96 – 108, doi: http://dx.doi.org/https://doi.org/10.1016/j.ress.2015.12.002.

[Sinou, 2009] Sinou, J.-J. (2009). *A Review of Damage Detection and Health Monitoring of Mechanical Systems from Changes in the Measurement of Linear and Non-linear Vibrations*, pages 643–702.

[Sofi and Romeo, 2016] Sofi, A. and Romeo, E. (2016). A novel interval finite element method based on the improved interval analysis. *Computer Methods in Applied Mechanics and Engineering*, 311:671 – 697, doi: http://dx.doi.org/https://doi.org/10.1016/j.cma.2016.09.009.

[Sui, 2017] Sui, L. (2017). *Uncertainty management in parameter identification*. PhD thesis.

[Sui et al., 2018] Sui, L., Feissel, P., and Denœux, T. (2018). Identification of elastic properties in the belief function framework. *International Journal of Approximate Reasoning*, 101:69 – 87.

[Tam et al., 2017] Tam, J. H., Ong, Z. C., Ismail, Z., Ang, B. C., and Khoo, S. Y. (2017). Identification of material properties of composite materials using nondestructive vibrational evaluation approaches: A review. *Mechanics of Advanced Materials and Structures*, 24(12):971–986, doi: http://dx.doi.org/10.1080/15376494.2016.1196798.

[Teughels and Roeck, 2005] Teughels, A. and Roeck, G. D. (2005). Damage detection and parameter identification by finite element model updating. *Revue Européenne de Génie Civil*, 9(1-2):109–158.

[Tikhonov, 1995] Tikhonov, A. N. (1995). *Numerical Methods for the Solution of Ill-Posed Problems*, volume 328. Springer.

[Tikhonov et al., 1995] Tikhonov, A. N., Goncharsky, A., Stepanov, V. V., and Yagola, A. (1995). Numerical methods for the solution of ill-posed problems.

[Tipping, 2004] Tipping, M. E. (2004). *Bayesian Inference: An Introduction to Principles and Practice in Machine Learning*, pages 41–62. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Toja-Silva et al., 2014] Toja-Silva, F., Favier, J., and Pinelli, A. (2014). Radial basis function (rbf)-based interpolation and spreading for the immersed boundary method. *Computers Fluids*, 105:66 – 75, doi: http://dx.doi.org/https://doi.org/10.1016/j.compfluid.2014.09.026.

[Tornil-Sin et al., 2010] Tornil-Sin, S., Puig, V., and Escobet, T. (2010). Set computations with subpavings in matlab: The scs toolbox. In *2010 IEEE International Symposium on Computer-Aided Control System Design*, pages 1403–1408.

[Truman and Terlaje, 2006] Truman, K. and Terlaje, G. (2006). *Damage Detection Using Static Response Data and Optimality Criterion*, volume 140, pages 685–696.

[van Ravenzwaaij et al., 2016] van Ravenzwaaij, D., Cassey, P., and Brown, S. (2016). A simple introduction to markov chain monte–carlo sampling. *Psychonomic Bulletin Review*, 25, doi: http://dx.doi.org/10.3758/s13423-016-1015-8.

[Walley, 1991] Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall.

[Wang and Hickernell, 2000] Wang, X. and Hickernell, F. (2000). Randomized halton sequences. *Mathematical and Computer Modelling*, 32(7):887 – 899, doi: http://dx.doi.org/https://doi.org/10.1016/S0895-7177(00)00178-3.

[Xiao, 2005] Xiao, N. (2005). *Interval finite element approach for inverse problems under uncertainty*. PhD thesis, Georgia Institute of Technology.

[Xiao et al., 2013] Xiao, N., Fedele, F., and Muhanna, R. (2013). Inverse problems under uncertainties – an interval solution for the beam finite element.

[Xiaoguang et al., 2019] Xiaoguang, W., Weiliang, H., and Linggong, Z. (2019). Novel interval parameter identification method using augmented fourier series-based polynomial surrogate model. *IEEE Access*, 7:70862–70875.

[Zadeh, 1965] Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338 – 353, doi: http://dx.doi.org/https://doi.org/10.1016/S0019-9958(65)90241-X.

[Zhang, 2005] Zhang, H. (2005). *Nondeterministic Linear Static Finite Element Analysis: An Interval Approach*. PhD thesis, Georgia Institute of Technology.

[Zio and Pedroni, 2013] Zio, E. and Pedroni, N. (2013). Literature review of methods for representing uncertainty. *Foundation for an Industrial Safety Culture, Toulouse, France (ISSN 2100-3874)*, (2013-03 of the Cahiers de la Sécurité Industrielle).